

Thomas Gruschwitz

Einführung
einer Business Intelligence Lösung
bei einem mittelständischen Unternehmen
der Solarindustrie

Hochschule Mittweida
University of Applied Sciences



Diplomarbeit

Erstprüfer

Prof. Dr. rer. pol. Silke Meyer

Zweitprüfer

Dipl.-Betriebswirt (BA) Roger Kunig

Mittweida, Juni 2010

Inhalt

I Abbildungsverzeichnis.....	III
II Tabellenverzeichnis.....	V
1 Einleitung.....	6
1.1 Ausgangssituation.....	7
1.2 Zielstellung.....	9
1.3 Aufbau der Arbeit.....	10
2 Business Intelligence.....	11
2.1 Data Warehousing.....	17
2.1.1 Begriff und Eigenschaften eines DWH.....	18
2.1.2 ETL.....	22
2.1.3 Metadaten eines DWH.....	30
2.2 Datenanalyse.....	32
2.2.1 OLAP.....	33
2.2.2 Data Mining.....	40
2.3 Portale.....	40
2.4 Standardsoftware und BI-Markt.....	42
2.5 Einführung einer BI-Lösung mit Sicherheitskonzept.....	46
3 Konzept und Implementierung einer BI-Lösung bei AVANCIS.....	51
3.1 Projektdefinition.....	53
3.2 Ist- und Sollanalyse.....	54
3.3 Design & Implementierung.....	60
3.3.1 Datenintegration.....	60
3.3.2 Analyse und Reporting.....	66
3.3.3 Bereitstellung.....	77
3.4 Produktivsetzung, Betrieb und Wartung.....	81
4 Fazit & Perspektive.....	83
4.1 Integration von Data Mining.....	88
III Literaturverzeichnis.....	I
IV Anlage A.....	VII
V Anlage B.....	IX

Inhalt	II
VI Anlage C.....	XIII
VII Selbständigkeitserklärung.....	XIV

I Abbildungsverzeichnis

Abbildung 1: Datenwachstum in der ORACLE-Datenbank innerhalb eines Jahres.....	8
Abbildung 2: Aufbau der Diplomarbeit.....	10
Abbildung 3: Ordnungsrahmen für BI-Systeme.....	14
Abbildung 4: BI Drei-Phasenmodell	15
Abbildung 5: Phasen der BI und deren Aktivitäten.....	16
Abbildung 6: Architekturvarianten DWH.....	20
Abbildung 7: Übersicht ETL-Prozess	23
Abbildung 8: Einordnung von syntaktischen und semantischen Mängeln in Mängelklassen.....	25
Abbildung 9: Filterung, Harmonisierung, Aggregation und Anreicherung.....	28
Abbildung 10: Dreidimensionales Datenmodell.....	38
Abbildung 11: Analysetechnik Slice.....	39
Abbildung 12: BI-Anbieterportfolio.....	45
Abbildung 13: Schema des erweiterten BI-Vorgehensmodells.....	48
Abbildung 14: Modell zur Berechtigungskonzeption.....	50
Abbildung 15: Entwurf zur Darstellung von prozesskritischen Parametern.....	57
Abbildung 16: Ausbuchung von Solarmodulen und deren Erfassung.....	59
Abbildung 17: GUI STAS CONTROL ETL.....	61
Abbildung 18: Prozessbeispiel in STAS CONTROL ETL.....	63
Abbildung 19: Schematischer Aufbau der Tabellen Ausbuchungen (EQ_VALUE und MOVEMENT) und Produktionsmengen (DAY_REPORT).....	65
Abbildung 20: Transformation der Tabelle DAY_REPORT.....	65
Abbildung 21: Aggregation der Ausbuchungsmengen auf die Ebene Datum, Equipment, Grund.....	66
Abbildung 22: logische Cognosarchitektur.	67
Abbildung 23: Cognos Business Intelligence Tools.....	69
Abbildung 24: GUI des Cognos 8 Report Studio.....	71
Abbildung 25: Deklarierte Beziehungen zwischen den Spalten der DWH-Tabellen.....	74
Abbildung 26: Modellierung eines OLAP-Würfels zur Darstellung der Ausbuchungswerte.....	74

Abbildung 27: Dimensionen des generierten OLAP-Würfels	75
Abbildung 28: Trendchart.....	76
Abbildung 29: Paretodiagramm.....	76
Abbildung 30: Implementierte BI-Struktur bei AVANCIS.....	83

II Tabellenverzeichnis

Tabelle 1: Differenzierung von Data Mart und DWH.....	21
Tabelle 2: Verteilung des BI-Marktes in Deutschland.....	46
Tabelle 3: Lizenztypen Cognos 8.4.....	78
Tabelle 4: Gegenüberstellung von vorgefertigter und individueller BI-Lösung.....	86

1 Einleitung

Die fetten Jahre sind vorbei. So titelte am 23. Januar 2010 *DIE WELT*. Während vor einigen Jahren das wirtschaftliche Umfeld den Unternehmen der Solarbranche eine hohe Gewinnmarge¹ für ihre Produkte zusicherte, stehen sie nun vor entscheidenden Herausforderungen. Die Kürzung staatlicher Subventionen für die Einspeisung von solarerzeugtem Strom sowie die auf den Markt drängende Konkurrenz aus Asien haben zu einem Preisverfall geführt, der deutschen Solarunternehmen zukunftssichernde Entscheidungen abverlangt.

Um diese Entscheidungen treffen zu können, ist es für das Management maßgeblich, über Daten und Fakten wie Produktionsentwicklung, Umsatz- und Absatzzahlen sowie Tendenzen zur Gewinnerzielung genauestens informiert zu sein. Damit die Entscheidungsträger in ihrem Urteils- und Entschlussfindungsprozess optimal und mit, für sie selbst möglichst geringem Aufwand unterstützt werden können, ist es erforderlich, Unternehmensdaten durch leistungsfähige Software stets und umfassend für den zuständigen und entscheidungsbetragten Personenkreis zur Verfügung zu stellen.

Die beschriebene wirtschaftlich schwierige Lage in der Solarindustrie erfordert Veranlassungen, welche sich auf konkrete und fundierte Daten und Entwicklungen stützen. So können etwa Effizienzlücken besser erkannt und geschlossen sowie das Verhältnis von Angebot und Nachfrage überlegter eingeschätzt werden. Entscheidungen, die vermehrt auf Intuition und Impulsen als auf konkreten Daten und Trendentwicklungen beruhen, sind in der derzeitigen wirtschaftlichen Lage meist unvorteilhaft. Vielmehr sollten sie auf realen Zahlen und Fakten fußen.

Auch das relativ junge Unternehmen *AVANCIS* bewegt sich in diesem Umfeld und steht vor der Herausforderung, Informationen auf Basis interner und externer Daten zu gewinnen und damit die Führungskräfte in die Lage zu versetzen, unangreifbare Entscheidungen zu treffen. Gegenstand dieser Diplomarbeit ist diesbezüglich, die aktuellen Techniken im Rahmen von *Business Intelligence* (BI) zu beleuchten, den Einführungsprozess zu begleiten und Einfluss auf diesen zu nehmen.

¹ Im Jahr 2006 betrug die Marge pro Produkteinheit (Verhältnis von Kosten und Erlös) etwa 30 Prozent; Vgl. [Sch10]

1.1 Ausgangssituation

Die *AVANCIS GmbH & Co. KG* ist ein Unternehmen des französischen Konzerns *Saint Gobain* mit Sitz im nordsächsischen Torgau. Seit dem Jahr 2008 bringt *AVANCIS* Dünnschicht-Solarmodule mit einer jährlichen Gesamtleistung von 20 MW auf den Markt, welche das Unternehmen derzeit ausschließlich am Standort in Torgau produziert.

Ein weiterer Standort befindet sich in München. An diesem wird die Optimierung der eingesetzten Technologie in der Forschungs- und Entwicklungsabteilung vorangetrieben. Des Weiteren ist hier die Fabrikplanung zu finden, welche die Aufgabe hat weitere Werke für die Herstellung von Solarmodulen zu konzeptionieren. Mit diesen Fabriken soll die Ausbringungsleistung jeweils um jährlich 100 MW erhöht werden.¹

AVANCIS fertigt die Dünnschichtsolarmodule auf Basis einer CIS-Technologie. *CIS* steht dabei für die Elemente Kupfer, Indium und Selen. Zur Absorption des Sonnenlichtes werden dazu die Elemente durch verschiedene Beschichtungsverfahren auf ein Glassubstrat aufgetragen, wodurch die Verwendung von Silizium umgangen wird.

Im Zuge der Start-Up-Phase wurde bei *AVANCIS* eine IT-Infrastruktur aufgebaut. Diese basiert u. a. auf dem Enterprise Resource Planning System *SAP ERP ECC 5.0* sowie mit *Siemens SILOC* auf einem System zur vertikalen Integration von Maschinensteuerungen und dem ERP-System. Durch diese operativen Systeme sieht sich das Unternehmen mit einer kontinuierlich wachsenden Datenbasis konfrontiert, so dass die Schwierigkeit wächst, aus den vorhandenen Daten die benötigten Informationen abzuleiten. Ein besonders rapides Wachstum des Datenbestandes ist dabei innerhalb der *ORACLE* Datenbank des Systems *Siemens SILOC* zu verzeichnen. Mit Hilfe von Datentelegrammen der Produktionseinheiten gelangen derzeit etwa 500 Daten pro Minute in die Datenbank, so dass innerhalb eines Jahres etwa 260.000.000 zusätzliche Daten verarbeitet werden müssen, wie die Abbildung 1 verdeutlicht.

¹ Am 08. Juni 2010 gab Saint Gobain bekannt, dass am Standort in Torgau eine weitere Fabrik zur Produktion von Solarmodulen errichtet und innerhalb des 1. Quartals im Jahr 2012 die Produktion aufnehmen wird.

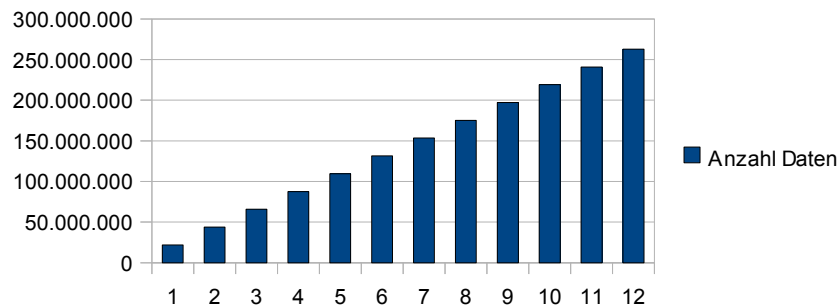


Abbildung 1: Datenwachstum in der ORACLE-Datenbank innerhalb eines Jahres

Ein weiteres Problem ist, dass Siemens *SILOC* nur rudimentäre Auswertungsmöglichkeiten von detaillierten Daten über die GUI bereitstellt. Innerhalb der Benutzerschnittstelle können verschiedene Filter gesetzt und anschließend das Ergebnis in eine CSV-Datei¹ exportiert werden. Diese Datei kann wiederum in *MS Excel*² importiert werden, so dass die Daten zur Auswertung zur Verfügung stehen. Dieser Prozess ist sehr zeitaufwändig und belastet bei Abfragen über größere Zeitintervalle das operative System.

Das SAP-Release *SAP ERP ECC 5.0* bietet verschiedene Auswertungsmöglichkeiten wie spezielle Transaktionen, Tabellenausgaben oder die Queryfunktionen. Das Erstellen von Queries setzt dabei Hintergrundwissen zur SAP-Tabellenstruktur, betriebswirtschaftliche Prozesskenntnis sowie eine gewisse Einarbeitungszeit voraus. Daher ergibt sich diesbezüglich eine geringe Flexibilität, so dass viele Daten aus dem ERP-System in Excel-Formate exportiert und innerhalb der Tabellenkalkulation ausgewertet werden. Darüber hinaus können mit Hilfe von Transaktionen wie der *SE16* SAP-Tabellen ausgegeben und deren Daten gefiltert, sortiert oder etwa aggregiert werden. Der Nachteil dieser Auswertungsmöglichkeit besteht darin, dass keine Verknüpfung von Tabellen möglich ist.

Innerhalb der SAP-Module stehen weiterhin für Abfragezwecke verschiedene Transaktionen zur Verfügung, welche etwa durch die Ergänzung von Spalten an den Informationsbedarf angepasst werden können. Diese Anpassung muss jedoch durch die Modifikation des Programmcodes der Transaktion in der SAP-Programmiersprache *ABAP*³

1 CSV steht in diesem Zusammenhang für Comma-Separated Values.

2 Der Softwarehersteller Microsoft verwendet innerhalb seiner Produktbezeichnungen die Abkürzung MS für Microsoft.

3 Advanced Business Application Programming

erfolgen. Diese Aufgaben können aufgrund der knappen Ressourcen nicht durch die interne AVANCIS-IT wahrgenommen werden, so dass die Umsetzung durch einen externen Dienstleister wie *aii* erfolgen muss und sich somit in der Regel kostenintensiv und unflexibel gestaltet.

Da die Auswertungsmöglichkeiten der vorliegenden operativen Systeme SAP ERP ECC 5.0 und *SILOC* unzureichend sind, erfolgt die Auswertung von Daten der operativen Systeme vordergründig in MS Excel. Auswertungen in Excel bieten die notwendige Flexibilität, setzen jedoch für komplexe Datenanalysen gewisse Kenntnisse voraus und bergen Schwierigkeiten bei der Verarbeitung großer Datenmengen¹ und der Verteilung².

1.2 Zielstellung

Nur durch die bedarfsorientierte Bereitstellung von Informationen kann das Management in die Lage versetzt werden, Entscheidungen auf der Basis von objektiven Kriterien zu treffen und so die Zukunft des Unternehmens zu gestalten. Die eingesetzten Systeme erfüllen ihre Aufgabe im operativen Geschäft, stoßen jedoch bei der entscheidungsunterstützenden Aufbereitung der Daten an ihre Grenzen.

Für die Entlastung der operativen Systeme bei Abfragen und Analysen ist daher die Überführung von Daten in eine separate Datenbank für Abfragezwecke zielführend und diese in einem für Auswertungen optimierten System aufzubereiten und zu präsentieren.

Die Diskrepanz zwischen der benannten und der angestrebten Situation lässt sich durch die Einführung einer weiteren Unternehmenssoftware beheben. Daher wurde bei *AVANCIS* ein Projekt zur Implementierung einer *Business Intelligence Lösung* initiiert.

1 Tabellenblätter in MS Excel können maximal 65536 Zeilen enthalten.

2 I. S. v. Schaffung eines zentralen Zugangs auf die Ergebnisse der Auswertungen in MS Excel

1.3 Aufbau der Arbeit

Das Thema *Einführung einer Business Intelligence Lösung bei einem mittelständischen Unternehmen der Solarindustrie* wird wie in Abbildung 4 dargestellt in vier Kapiteln behandelt.

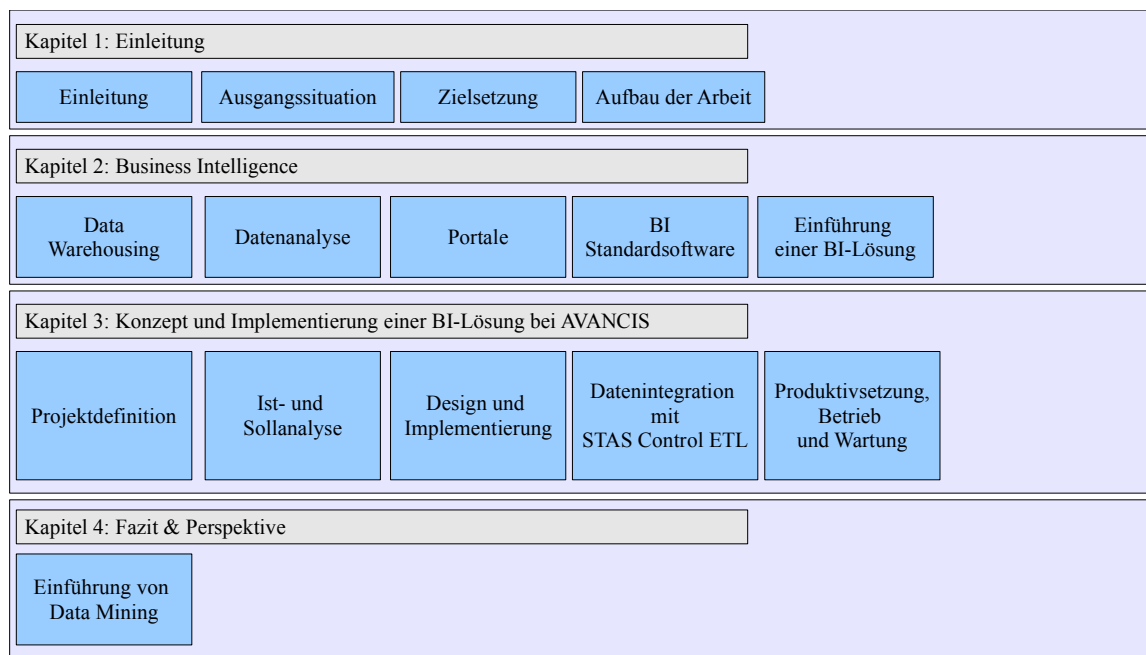


Abbildung 2: Aufbau der Diplomarbeit

In Kapitel 2 werden die notwendigen Begriffe *Business Intelligence*, *Data Warehousing*, *BI-Portale* und *Standardsoftware* definiert. Es werden zudem die Schritte zur Einführung einer *BI-Lösung* aufgezeigt.

Darauf aufbauend wird im nächsten Kapitel die praktische Umsetzung der in Kapitel 2 genannten theoretischen Vorgehensweisen am Beispiel der Einführung einer *Business Intelligence-Lösung* bei *AVANCIS* erläutert. Besondere Beachtung finden dabei die Applikationen *STAS CONTROL ETL powered by Talend* und *IBM Cognos 8.4*.

In Kapitel 4 wird das Projekt zusammengefasst und mögliche Perspektiven zur Weiterentwicklung der *BI-Lösung* diskutiert.

2 Business Intelligence

Mittels *Business Intelligence* werden Führungskräfte des operativen, taktischen und strategischen Managements durch die bedarfsorientierte Bereitstellung von entscheidungsrelevanten Daten im Zuge der Entscheidungsfindung unterstützt. Dazu wird „operatives Datenmaterial zur Informations- und letztlich zur Wissensgenerierung“¹ aufbereitet und abgespeichert sowie Auswertungs- und Präsentationsfunktionalität angeboten.² *Business Intelligence* kann dabei als der aktuelle Stand einer dynamischen Entwicklung angesehen werden, die seit den 1960er Jahren mit dem Aufkommen der ersten *Management Informationssysteme* (MIS) begann.³

In der breiten Wissenschaft wurde in den 1990er Jahren die Auffassung vertreten, dass, im Zuge der erweiterten technischen Möglichkeiten, unter dem Sammelbegriff *Management Support Systeme* die Gesamtheit der managementunterstützenden IT-Systeme erfasst wurden.⁴ Dennoch entstand in der Praxis mit *Business Intelligence* ein weiterer Begriff, der in der frühen marketingorientierten Phase als Sammelbezeichnung für Frontendwerkzeuge zur Analyse Verwendung fand.⁵

Erstmals benutzt wurde *Business Intelligence* im Jahr 1958 von *Luhn*, der zu dieser Zeit für den amerikanischen Informationstechnikkonzern IBM⁶ tätig war. *Howard Dresner*, ein Analyst der *Gartner Group*⁷, knüpfte 1996 an *Luhns* Vorstellungen an und präzierte diese.⁸

„By 2000, Information Democracy will emerge in forward-thinking enterprises, with Business Intelligence information and applications available broadly to employees, consultants, customers, suppliers, and the public. The key to thriving in a competitive marketplace is staying ahead of the competition. Making sound business decisions based on accurate and current information takes more than intuition. Data analysis, reporting, and

1 [Cha04, S. 120]

2 Vgl. ebenda, S. 120

3 Im Anlage A befindet sich eine geschichtliche Abhandlung zum Thema BI.

4 Vgl. [Glu08, S. 89]; [Kem04, S. 2]

5 Vgl. [Kem04, S. 2]

6 International Business Machines Corporation

7 Das heutige Unternehmen Gartner (seit 2001) wurde 1979 am Hauptsitz in Stamford, USA gegründet.

8 Vgl. [Zim09]

query tools can help business users wade through a sea of data to synthesize valuable information from it - today these tools collectively fall into a category called - Business Intelligence.“¹

Nach einem anfänglichen Akzeptanzproblem des Begriffes konnte sich *Business Intelligence* im weiteren Verlauf in der Praxis durchsetzen, da führende Hersteller für Analysewerkzeuge alle angebotenen Produkte unter dem Sammelbegriff BI zusammenfassten.² Durch den Einzug in die Praxis wurde der Begriff zeitversetzt auch in der Wissenschaft angewandt und kontrovers diskutiert³.

Die wissenschaftliche Diskussion führte nicht zu einer klaren Begriffsabgrenzung und -definition. *Peter Mertens* machte 2002 bei der Analyse von gängigen BI-Definitionen auf diesen Missstand aufmerksam, in dem er auf sieben unterschiedliche Definitionsvarianten in seiner Publikation „Simulation in der strategischen Unternehmensplanung“ einging:

- „BI als Fortsetzung der Daten- und Informationsverarbeitung: Informationsverarbeitung für die Unternehmensleitung
- BI als Filter in der Informationsflut
- BI = MIS, aber besonders schnelle/flexible Anwendungen
- BI als Frühwarnsystem („Alerting“)
- BI = Data Warehouse
- BI als Informations- und Wissensspeicherung
- BI als Prozess: Symptombehebung → Diagnose → Therapie → Prognose → Therapiekontrolle“⁴

Alle Definitionen weisen bei genauerer Betrachtung eine Gemeinsamkeit auf, denn sie enthalten Konzepte bzw. Technologien⁵ die sich ganz oder teilweise *Business Intelligence* zuordnen lassen. „Ein Grundkonsens besteht dabei darin, dass die Techniken und Anwendungen des Business Intelligence entscheidungsunterstützenden Charakter

1 [Tec10]

2 Vgl. [Kem07, S. 58]

3 In Teilen der Wissenschaft findet der Begriff MSS noch heute Verwendung und wird als Sammelbegriff für IT-basierte Managementunterstützung genutzt. Vgl. [Kemp04, S. 2]

4 [Mer02, S. 4]

5 Konzepte und Technologien im Sinne von Systemen

aufweisen sowie zur besseren Einsicht in das eigene Geschäft und damit zum besseren Verständnis in die Mechanismen relevanter Wirkungsketten führen sollen“⁶. *Chamoni* und *Gluchowski* unterstreichen diesen Gedanken und definieren BI als Sammelbegriff „zur Kennzeichnung von Systemen [...], die auf Basis interner Leistungs- und Abrechnungsdaten sowie externer Marktdaten in der Lage sind, das Management in seiner planenden, steuernden und koordinierten Tätigkeit zu unterstützen“⁷

Anhand dieser Definitionen lässt sich *Business Intelligence* einer Vielzahl von Systemen zuordnen, so dass die Anzahl dieser durch *Gluchowski* mit Hilfe eines Ordnungsrahmens begrenzt wurde. Er betrachtet die Systeme, wie in Abbildung 3 aufgezeigt wird, dabei von zwei verschiedenen Standpunkten: der Prozessphase, dargestellt auf der Y-Achse, und dem Schwerpunkt, dargestellt auf der X-Achse. Die Prozessphase verläuft dabei von der Datenbereitstellung zur Datenauswertung, der Schwerpunkt hingegen unterscheidet zwischen technik- und anwendungsorientierten Systemen.

6 [Glu08, S. 90]

7 [Cha04, S. 119]

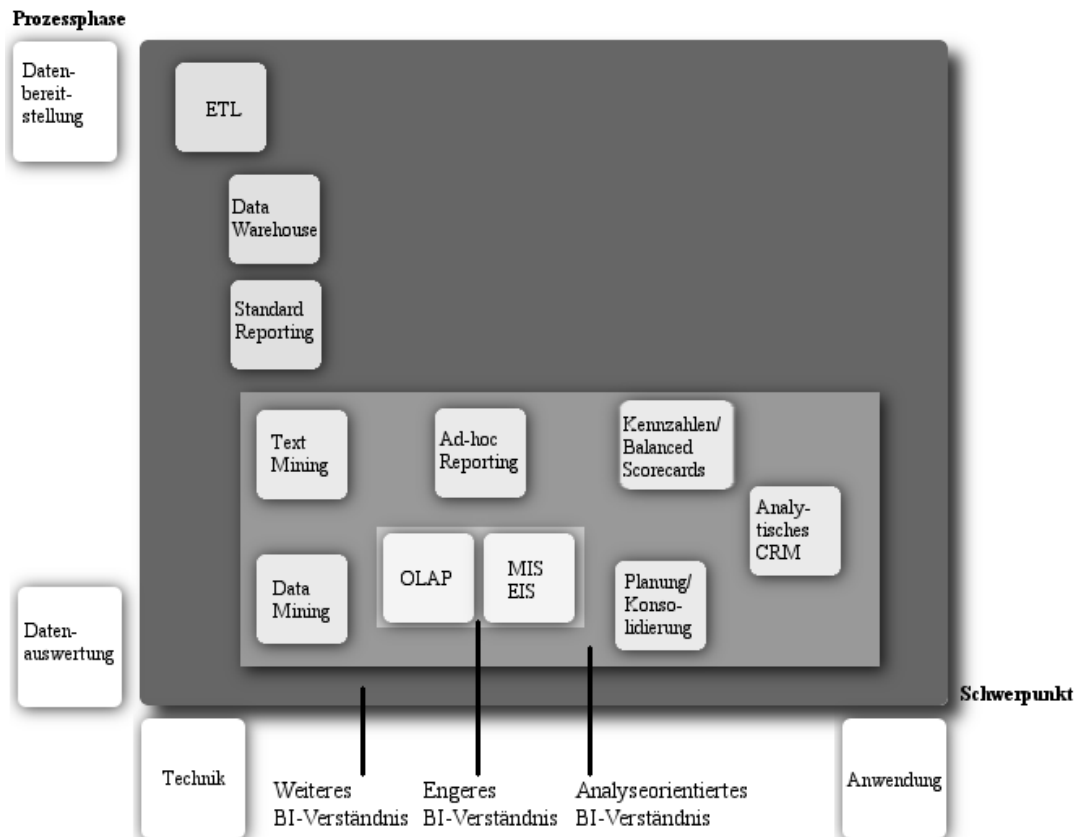


Abbildung 3: Ordnungsrahmen für BI-Systeme [Glu01, S.7]

Aus der Positionierung der Systeme innerhalb des Ordnungsrahmens leitet *Gluchowski* eine Zuordnung in ein engeres, analyseorientiertes und weiteres BI-Verständnis ab. Dem engeren BI-Verständnis werden dabei nur wenige Kernapplikationen wie Online Analytical Processing, Management Information Systems und Executive Information Systems zugeordnet, welche zur unmittelbaren Entscheidungsfindung beitragen.¹

Die Produkte und Anwendungen des engeren BI-Verständnisses werden im Rahmen des analyseorientierten BI-Verständnisses durch Komponenten ergänzt, welche modell- und methodenbasiert eine zielgerichtete Analyse von vorhandenen Daten ermöglichen.² Diese Eigenschaft trifft insbesondere auf Produkte zum Data und Text Mining sowie auf Generatoren zur Erstellung von Berichten auf multidimensionaler Datenbasis zu. Daran

¹ Vgl. [Kem04, S. 3]; [Glu08, S. 90]

² Vgl. [Glu08, S. 90]

schließen sich ganz oder teilweise konkrete Anwendungen an, die den analytisch arbeitenden Anwender unterstützen, wie Balanced Scorecard- und Kennzahlensysteme sowie Systeme zur Planung, Budgetierung und Konsolidierung sowie zum Risikomanagement.¹

Während *Gluchowski* eine Einordnung der BI-Systeme in einen zweidimensionalen Ordnungsrahmen vornimmt, legen *Grothe und Gentsch* den Schwerpunkt auf eine prozessorientierte Definition. Daher sehen sie *Business Intelligence* als Prozess, „welcher aus fragmentierten und inhomogenen unternehmensinternen und -externen Daten Wissen über die eigenen und über fremde Positionen, Potenziale und Perspektiven generiert“². Die Art und Weise der Prozesskettengestaltung richtet sich dabei nach dem Aspekt der Suche, wonach „keine vorgefertigten Meinungen bestätigt, sondern neue Erkenntnisse adaptiert sowie Hypothesen bestätigt oder abgelehnt werden“³. Unter Beachtung dieses Aspektes besteht der *Business Intelligence Prozess* nach *Grothe und Gentsch* aus den drei Phasen der Bereitstellung, Entdeckung und Kommunikation von entscheidungsrelevanten Informationen.⁴

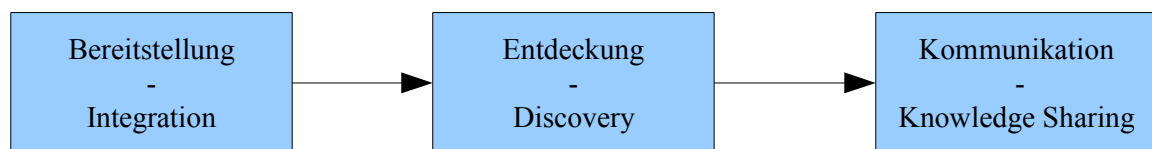


Abbildung 4: BI Drei-Phasenmodell

Zur Bewältigung dieser Prozessphasen sind *Business Intelligence Applikationen* notwendig, welche diesen Phasen wie folgt zugeordnet werden können.

Bereitstellung – Integration. In diesem Abschnitt werden relevante Quelldaten klassifiziert, in einem ETL-Prozess⁵ entscheidungsrelevant aufbereitet und anschließend in einer für die Analyse geeigneten Datenbank, dem Data Warehouse (DWH), bereitgehalten.

Entdeckung – Discovery. Dieser Phase werden Technologien zugeordnet, welche Berichte

1 Vgl. [Glu08, S. 90-91]

2 [Gen10]

3 [Teg05, S. 75]

4 Vgl. [Gro00, S. 20]

5 ETL ist die Kurzform für die Begriffe Extract, Transform, Load.

und Ad-hoc-Abfragen auf der Datenbasis des Data Warehouses erzeugen. Applikationen wie OLAP (Online Analytical Processing) und Data Mining ergänzen diese Phase und bieten die Möglichkeit zur Durchführung von komplexen Analysen.

Kommunikation - Knowledge Sharing. Über ein Webportal werden die Ergebnisse von Analysen den Anwendern zugänglich gemacht. Ziel ist es, anschließend auf Basis der gewonnenen Erkenntnisse Prozesse im Unternehmen zu überprüfen und anzupassen.

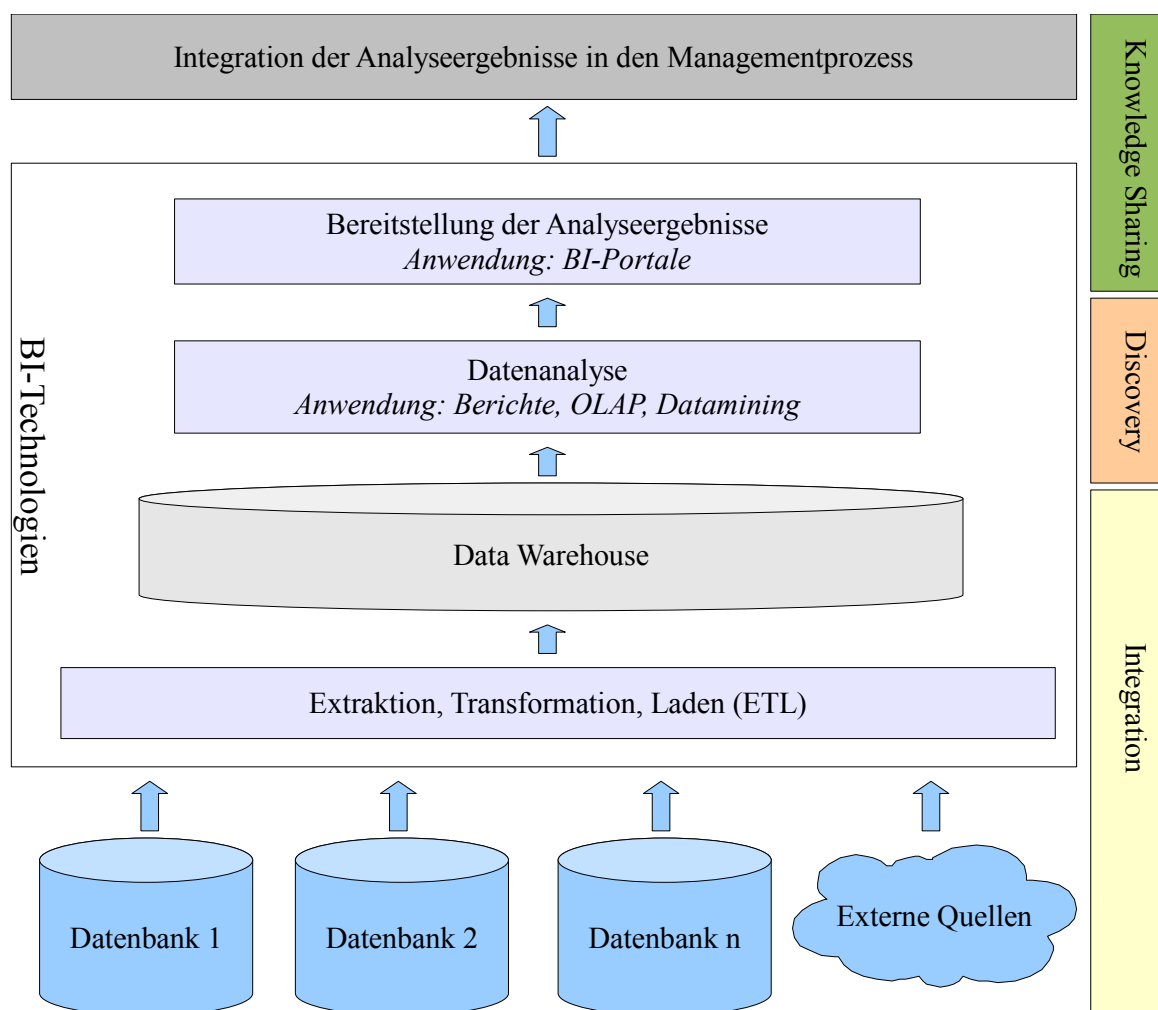


Abbildung 5: Phasen der BI und deren Aktivitäten. Nach [Gen09]

Zusammenfassend lassen sich dem *Business Intelligence* im weiteren Sinne alle direkt und indirekt für die Entscheidungsunterstützung eingesetzten Anwendungen zuweisen, die operatives Datenmaterial zur Informations- und Wissensgenerierung aufbereiten und

speichern sowie Auswertungs- und Präsentationsfunktionalität anbieten.¹ Das schließt auch Komponenten zur Datenbereitstellung wie das *Data Warehouse*, Werkzeuge zum Transformationsprozess der Daten sowie sämtliche analytische Anwendungen ein, die interdependent entlang einer Prozesskette von Bereitstellung, Entdeckung und Kommunikation mit dem Ziel der Managementunterstützung agieren.

Grundlage für die Entscheidungsunterstützung durch *Business Intelligence* bilden dabei u. a. die Daten der operativen Systeme sowie Daten externer Quellen, wobei die operativen Systeme nach *Gluchowski et al.* „gezielt die wertschöpfenden Geschäftsprozesse innerhalb des Unternehmens“² unterstützen. Transaktionsorientierte Daten sind nach *Bauer* und *Günzel* „meist zeitaktuell, nicht abgeleitet, autonom, aus einer Datenquelle und dynamisch.“³ Die Dynamik leitet sich daraus ab, dass operative Daten von vielen Schreibvorgängen und ständigen Modifikationen betroffen sind.⁴ Systeme, welche das Steuern der Prozesse ermöglichen sollen⁵, erfordern hingegen Daten, die „konsolidiert, integriert, stabil und meist aggregiert sind“⁶ Die Bereitstellung derartiger Daten sind Gegenstand des nächsten Abschnittes.

2.1 Data Warehousing

Eine wesentliche Aufgabe des *Business Intelligence* ist es, einen geeigneten und stimmigen Datenbestand zu Informations- und Analysezwecken bereitzustellen. Voraussetzung dafür ist eine entscheidungsorientierte Datenbasis, deren kontinuierliche Befüllung mit bereits existierenden und zukünftig generierten Daten gewährleistet sein muss.⁷ Data Warehouse, ETL-Prozess und Metadaten stellen Realisierungs- und Grundkonzepte dar und lassen sich dem Data Warehousing zuordnen.⁸ Sie werden in den folgenden Abschnitten thematisiert.

1 Vgl. [Kem04, S. 4]; [Glu08, S. 91]

2 [Cha04, S. 114]

3 [Bau04, S. 10]

4 Vgl. ebenda, S. 10

5 Vgl. [Gra04, S. 82]

6 [Bau04, S. 10]

7 Vgl. [Glu08, S. 117]

8 Vgl. ebenda, S. 117

2.1.1 Begriff und Eigenschaften eines DWH

Die ersten Ansätze zur Managementunterstützung in 1960er Jahren scheiterten nicht zuletzt daran, dass Management Information Systeme einen direkten Zugriff auf die operativen Systeme zur Datenbeschaffung nutzten. Dieses Problem wurde bei der Entwicklung des DWH-Konzeptes aufgegriffen, so dass als Leitgedanke und Lösungsansatz die konsequente Trennung von operativen und dispositiven Datenbeständen gilt. Die Datenhaltung nach dem DWH-Konzept verfolgt einen zentralistischen Ansatz, so dass die Bereitstellung der Daten auf „einer konsistenten [und] unternehmensweiten Datenhaltung“¹ basiert. Es „fordert [weiterhin] den Aufbau einer Datenbasis mit entscheidungsrelevanten Inhalten zur Unterstützung dispositiver Aufgaben“².

Zur inhaltlichen Ausrichtung des *Data Warehouse* hat insbesondere *Inmon* beigetragen, welcher als der Gründer des Data Warehouse Konzeptes gilt und den Begriff wie folgt definierte³:

„A data warehouse is a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management’s decision-making process.“

Die inhaltliche Ausrichtung eines Data Warehouse unterscheidet sich also wesentlich von der Ausrichtung operativer Datenhaltungssysteme. Sie ist geprägt durch vier wesentliche Eigenschaften: Subjektorientierung, Integration, Nicht-Volatilität und Zeitraumbezug.

Subjektorientierung. Das DWH grenzt sich unter anderem zu operativen Systemen in dem Punkt ab, dass die Daten nicht dazu dienen, Prozesse entlang der Wertschöpfungskette abzubilden. Vielmehr dient es dazu, entscheidungsrelevante bzw. dispositive Daten bedarfsorientiert zur Verfügung zu stellen.

Integration. In einem Data Warehouse werden Daten verschiedenster Datenquellen vereinigt. Die Daten müssen dabei eine Entscheidungsrelevanz aufweisen und können aus externen sowie aus unternehmensinternen Datenquellen stammen.

1 [Glu08, S. 117]

2 Ebenda, S. 118

3 [Clo10]

Nicht-Volatilität. Bewegungsdaten in operativen Systemen unterliegen einer ständigen Veränderung durch die auf den Daten ausgeführten Transaktionen. Im Gegensatz dazu sind die Daten in einem DWH für die langfristige Speicherung vorgesehen. Grundsätzlich werden Daten, welche einmal in einem DWH abgelegt wurden, nicht mehr verändert oder gelöscht. Bei einem großen Datenaufkommen wird es jedoch notwendig, ein Konzept für den Umgang mit Daten, welche ein gewisses Alter überschritten haben, zu erarbeiten.

Zeitraumbezug. Die Daten in operativen Systemen werden in der Regel zeitpunktbezogen abgespeichert. Die Daten eines DWH dienen hingegen zur Repräsentation einer Periode und werden daher mit Zeitraumbezug abgespeichert. Dieser Punkt wird jedoch häufig kritisiert, da aufgrund der Masse an Daten und zur Erzeugung von mehrdimensionalen Datenmodellen, häufig die unterste Zeitstufe abgelegt wird und die Daten für eine größere Zeitspanne zu verdichten sind.

Für *Inmon* ergaben sich aus der inhaltlichen Ausrichtung die Aufgaben eines Data Warehouses, welches themenorientierte¹ und integrierte² Informationen über lange Zeiträume und mit Zeitbezug zur Unterstützung von Entscheidungsträgern aus unterschiedlichen Quellen periodisch sammelt, aufbereitet und bedarfsgerecht zur Verfügung stellt.³

Neben der inhaltlichen Ausrichtung sind Aufbau und Gestaltung des Data Warehouse zu thematisieren. Die Architektur eines Data Warehouses beschreibt diesbezüglich die Anordnung der technologischen Konzepte in Bezug auf ihre Art und funktionellen Eigenschaften sowie die Verknüpfung der Elemente untereinander⁴. Innerhalb des DWH-Konzeptes wird dabei eine Unterscheidung zwischen zentralem und dezentralem Ansatz vorgenommen. Mit der sogenannten „*Hub-and-Spoke*“-Architektur wird häufig im Unternehmen die Verschmelzung des zentralen und dezentralen Ansatzes angewandt.

1 i. S. v. entscheidungsrelevant

2 i. S. v. vereinheitlicht

3 Vgl. [Inm96, S. 29-39]

4 Vgl. [Win10]; [Kem04, S. 23]

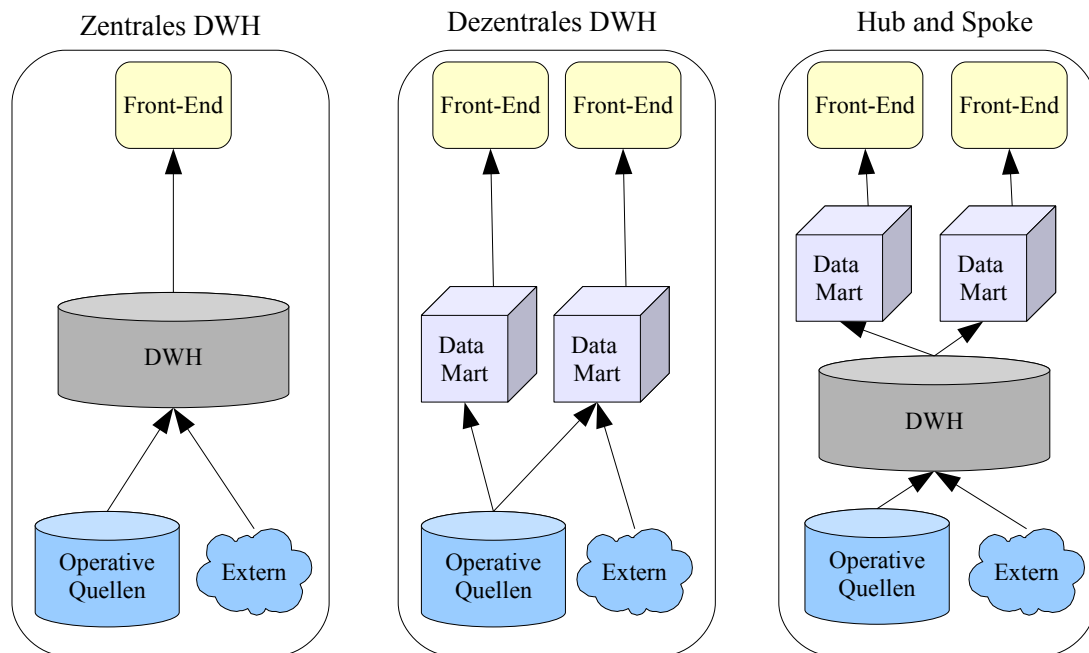


Abbildung 6: Architekturvarianten DWH. Nach [Abt10, S. 73]

In einem zentralen *Data Warehouse* werden „alle dispositiven Daten logisch zentral gespeichert und unter der Kontrolle eines einzigen Datenbankmanagementsystems verwaltet“¹. Dieses *Core Data Warehouse* ist eine nach dem DWH-Gedanken von den operativen Systemen isolierte Datenbank, welche eine unternehmensweit einheitliche Datenbasis zur Verfügung stellt. Damit besteht die Möglichkeit, logisch zusammengehörige aber physisch verteilte Daten unter der Verwendung eines Datenbankmanagementsystems in einem Core Data Warehouse zusammenzuführen².

Das Konzept des *zentralen Data Warehouses* zeigt sich bei wachsender Nutzerzahl sowie wachsenden Datenbeständen und die damit einhergehende Erhöhung der Abfragezeit in vielen Fällen als ungeeignet. Um die Abfragezeit niedrig zu halten wurde eine Verteilung der Verarbeitungs- und Administrationslast angestrebt. Dieser Ansatz wurde konsequent verfolgt, so dass inzwischen aus organisatorischen und technischen Gründen in vielen Unternehmen eine Abkehr vom zentralen Ansatz zu verzeichnen ist.³ Die Verteilung der dispositiven Daten erfolgt dabei in themen- bzw. abteilungsspezifische *Data Marts*.

1 [Kem04, S. 19]

2 Vgl. ebenda, S. 19

3 Vgl. [Bau04, S. 59-60]

Dadurch ist es möglich einzelnen Abteilungen bzw. Themengebieten zugeschnittene, performante¹ und detaillierte Analysen zur Verfügung zu stellen.

Die wesentlichen Unterscheidungsmerkmale zwischen *Data Marts* und einem *zentralen Data Warehouse* sind in der folgenden Tabelle aufgeführt.

Merkmale	Data Mart	Data Warehouse
Philosophie	anwendungsorientiert	anwendungsneutral
Adressat der Datenbereitstellung	Abteilung	Unternehmen
Granularität (Detaillierungsgrad der gespeicherten Daten)	Niedrig	Hoch
Datenmenge	Niedrig	Hoch
Menge historischer Daten	Niedrig	Hoch
Optimierungsziel	Abfragegeschwindigkeit	Datenmenge
Anzahl	mehrere	i. d. R. eins
Datenmodell	je nach Data Mart verschieden	Unternehmensweit

Tabelle 1: Differenzierung von Data Mart und DWH. [Abt10, S.73]

Um entscheidungsrelevante Daten in hoher Qualität und kostenoptimal anbieten zu können, hat sich eine Mischform aus zentralem und dezentralem Ansatz als Best Practice² erwiesen.³ Die Anordnung der Komponenten Data Warehouse und Data Marts erfolgt dabei so, dass für diese Architektur der Name *Hub-and-Spoke* bzw. der deutsche Begriff Nabe-Speiche verwendet wird. Dabei setzen als Speichen *Data Marts* auf der Nabe des zentralen Data Warehouse auf und speichern Teilextrakte des Gesamtdatenbestandes nochmals separat in physischer Form ab. Vorteil dieser Architektur ist die „Kombination zwischen weitgehender Integration und Vereinheitlichung der Daten in Verbindung mit der Flexibilität, die Daten auch örtlich anwendungsgerecht zur Verfügung zu stellen.“⁴

Grundlage für die Entwicklung der *Hub-and-Spoke Architektur* waren Probleme in der

1 Bei Unternehmen mit mehreren Standorten kann durch die physische Verteilung der Datenbestände auf die einzelnen Standorte eine bessere Performance erreicht werden.

2 Vorbildliche und nachahmenswerte Verfahrensweisen

3 Vgl. [Cha04, S. 65]

4 ebenda, S. 65

Praxis bei der Entwicklung eines unternehmensweiten Datenmodells. Für die Entwicklung müssen Fach- und Führungskräfte einzelner Abteilungen über einen längeren Zeitraum interdisziplinär zusammenarbeiten, so dass die Akzeptanz dieser Projekte im weiteren Verlauf abnahm. Daher entwickelten Abteilungen spezifische Data Marts, die sich jedoch als Insellösungen ohne die nötige unternehmensweite Integration erwiesen.

Mit der *zentralen*, der *dezentralen* und der Mischform *Hub-and-Spoke* können Unternehmen auf drei verschiedene Architekturkonzepte zur Implementierung einer Data Warehouse-Lösung zurückgreifen. Dabei kann keine der vorhandenen Lösungsmöglichkeiten als grundsätzlich optimal angesehen werden, so dass die Entscheidung für eine der Varianten auf Basis der vorhandenen Strukturen und der angestrebten Ergebnisse getroffen werden muss.

2.1.2 ETL

Daten, die in *OLTP¹-Systemen* hinterlegt sind, können aufgrund mangelnder Eignung nur teilweise direkt zur Managementunterstützung herangezogen werden. Diese bedingte Anwendbarkeit ist das Resultat der Ausrichtung von operativen Daten. Sie sind „auf die Anforderungen der mengen- und wertorientierten Abrechnungs-, Administrations- und Dispositionssysteme ausgerichtet und gewährleisten einen reibungslosen Ablauf des Tagesgeschäftes.“² Anwendungen zu Informations- und Analysezwecken stellen hingegen besondere Ansprüche an die Quantität und Qualität der Daten. Um diesen Ansprüchen gerecht zu werden, durchlaufen die Daten, welche aus operativen Systemen oder auch aus externen Quellen stammen, einen Transformationsprozess, in dem die Daten mit dem Ziel der Unterstützung des Managements aufbereitet werden.³

1 Online-Transaction-Processing

2 [Cham04, S. 115]

3 Vgl. [Kem04, S. 23]

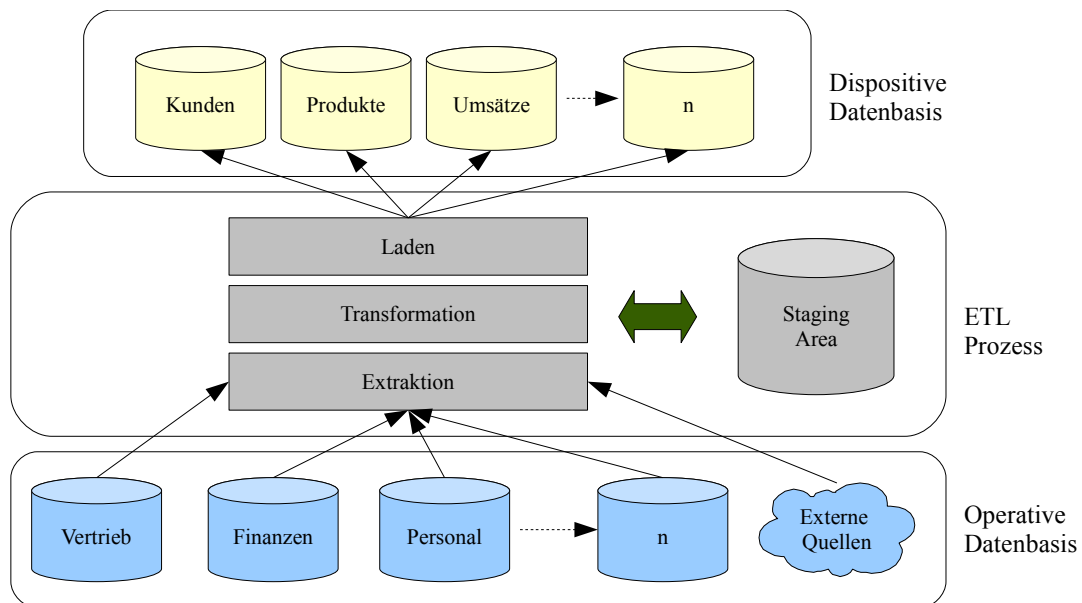


Abbildung 7: Übersicht ETL-Prozess

Der *Transformationsprozess*, auch *ETL-Prozess* (Extract, Transform, Load), umfasst alle Aktivitäten zur Überführung der operativen Daten in dispositive Daten. Dieser lässt sich in vier aufeinanderfolgende Teilprozesse untergliedern. Dazu gehört die Filterung, Harmonisierung, Aggregation und Anreicherung.¹

Filterung. Die Filterung umfasst die Extraktion der operativen Daten sowie die Bereinigung dieser Daten von syntaktischen und inhaltlichen Defekten.

Harmonisierung. In der Prozessphase der Harmonisierung werden betriebswirtschaftliche Daten abgestimmt.

Aggregation. Unter Aggregation wird die Verdichtung der Daten zum Beispiel über einen definierten Zeitraum verstanden.

Anreicherung. Im Anschluss an die Filterung, Harmonisierung und Aggregation werden betriebswirtschaftliche Kennzahlen gebildet und gespeichert.

Innerhalb des ersten Teilprozesses Filterung werden externe und operative Daten extrahiert und bereinigt. Die Extrahierung der Daten erfolgt ohne Modifikation bezüglich Format und

¹ Vgl. [Kem04, S. 24]

Inhalt in einen Arbeitsbereich des DWH. In diesen sogenannten *Staging Areas* werden die Daten für die anschließende Bereinigung von syntaktischen und semantischen Mängeln bereitgehalten. Die syntaktischen wie auch semantischen Mängel lassen sich dabei hinsichtlich Erkennung und Behebung drei verschiedenen Mängelklassen zuordnen.

Den Mängeln erster Klasse werden Defekte zugeordnet, die automatisiert erkannt und auch auf diese Weise behoben werden können. Voraussetzung dafür ist, dass die Datenmängel vor der Extraktion aus den operativen Systemen bekannt sind und daher mit Hilfe einer Routine eliminiert werden können. Zu dieser Klasse lassen sich syntaktische Mängel zuordnen, die in den operativen Systemen aufgrund der Verwendung von Format-, Steuer- oder Sonderzeichen zur Dokumentation beispielsweise von Stornobuchungen entstanden sind. Mit Hilfe von implementierten *Mapping-Tabellen* können die syntaktischen Mängel automatisch erkannt und behoben werden¹.

Semantische Mängel der ersten Mängelklasse zeichnen sich durch fehlende Daten wie etwa nicht vorhandene Ist-Werte in den Tabellen der operativen Systeme aus, welche innerhalb der Extraktionsphase identifiziert und nach definierten Regeln durch äquivalente Werte ersetzt werden. Dazu gehören etwa Planwerte des Monats, Ist-Werte des Vormonats oder Ist-Werte des Vorjahresmonats.²

Während Mängel erster Klasse automatisch erkannt und korrigiert werden, muss die Korrektur von Mängeln zweiter Klasse manuell durch Fachkräfte mit informationstechnischem und betriebswirtschaftlichem Hintergrund erfolgen. Auch Mängel der zweiten Klasse können semantische oder syntaktische Defekte aufweisen. *Chamoni* und *Gluchowski* verweisen darauf, dass „bei den syntaktischen Mängeln [...] hierbei beispielsweise bislang unberücksichtigte Syntaxvarianten der operativen Datenquellen entdeckt werden [können], die [...] manuell in den Extrakten berichtigt werden müssen.“³

Semantische Mängel der zweiten Klasse können durch Plausibilitätskontrollen, dazu zählen vor allem Vergleiche von Bilanz- oder Kontrollsummen, grundlegende Wertebereichsüberprüfungen oder Mustererkennungsverfahren auf Basis des Data Mining⁴

1 Vgl. [Cha04, S. 120]

2 Vgl. [Kem04, S. 26]

3 [Cha04, S. 121]

4 Vgl. [Kem04, S. 26-27]

erkannt werden. Die Ursache dieser Fehler liegt meist im operativen Systemumfeld. Daraus ergeben sich grundsätzlich zwei verschiedene Lösungsansätze. Eine kurzfristige Lösung verspricht die manuelle Modifikation der Daten in der Filterungsschicht des Data Warehouses. Dennoch sollte die Behebung der semantischen Mängel in den operativen Quellsystemen angestrebt werden, um manuellen Korrekturaufwand weitgehend auszuschließen.

Den Mängeln dritter Klasse werden Defekte zugeordnet, die weder automatisiert erkannt, noch automatisiert behoben werden können. In diesem Umfeld sind ausschließlich semantische Defekte zu beheben, da die Datensyntax stets vollständig beschrieben wird. Syntaxfehler können somit immer automatisch identifiziert werden. Anders verhält sich dies bei semantischen Mängeln, deren Erkennung sowie Behebung in vielen Fällen manuell erfolgen muss. Die fachlich nicht korrekten Daten zeichnen sich durch Mängel aus, die nur durch betriebswirtschaftliche Fachkräfte erkannt und behoben werden können. Aufgrund der Übereinstimmung der Ursachen erfolgt die Behebung der Defekte analog zur zweiten Mängelklasse. Als Lösungsansätze stehen die Modifikation der Daten innerhalb der Filterungsschicht und die Behebung der Fehler in den Quellsystemen zur Verfügung.

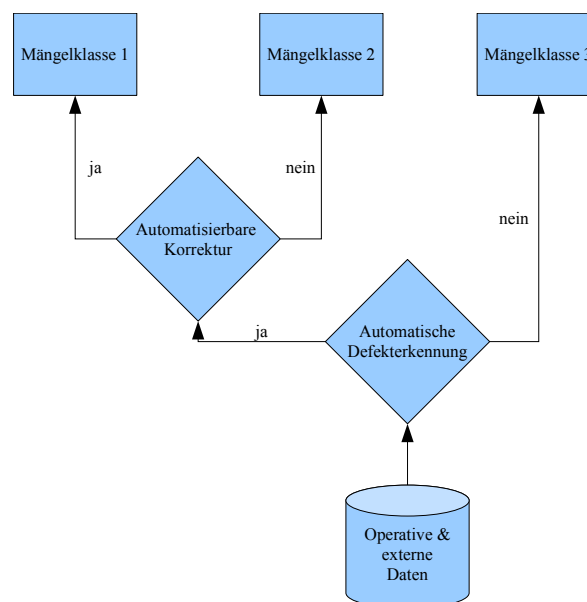


Abbildung 8: Einordnung von syntaktischen und semantischen Mängeln in Mängelklassen. Nach [Wil04, S. 41]

Nach der Extraktion und der anschließenden Bereinigung liegen in den Arbeitsbereichen des DWH bereits dispositiv verwendbare Daten mit geringer Granularität vor. Um entscheidungsrelevante Informationen performant dem adressierten Nutzerkreis zur Verfügung zu stellen, muss eine themenbezogene Gruppierung der Daten angestrebt werden. Dabei steht den Unternehmen zur physischen Zusammenführung inzwischen ein breites Spektrum an ETL-Werkzeugen zur Verfügung. Im Vorfeld dieser Zusammenführung von Datensätzen ist zu prüfen, ob eine syntaktische Harmonisierung bezüglich Schlüsseldisharmonien, unterschiedlich codierten Daten, Synonymen wie auch Homonymen notwendig wird. Die Vorgehensweise zur syntaktischen Harmonisierung kann in der Anlage C nachgelesen werden.

Die Verwendung von ETL-Werkzeugen hat zur Folge, dass die Herausforderung vordergründig in der betriebswirtschaftlichen Abgleichung der Daten liegt. Diese Aufgabe wird durch die Implementierung von Transformationsregeln realisiert, die eine hohe betriebswirtschaftliche Kompetenz des für dieses Teilgebiet Verantwortlichen voraussetzt.

Als weiterer Teilschritt innerhalb der betriebswirtschaftlichen Harmonisierung wird die Überführung detaillierter Daten in die gewünschte Granularität der dispositiven Daten angesehen. Dazu wird das „operative Datenmaterial in Bezug auf die betriebswirtschaftliche Bedeutung, die gebiets- und resortspezifische Gültigkeit, die Währung oder die Periodenzuordnung“¹ mittels Transformationsregeln in einheitliche Werte überführt.

Nach Abschluss der Harmonisierung liegt ein bereinigter konsistenter Datenbestand auf Granularitätsebene vor.

In der dritten Transformationsphase *Aggregation* werden die harmonisierten, gefilterten Daten der angestrebten Granularität um Verdichtungsstrukturen erweitert. „Zu diesem Zweck werden in der Regel eine Reihe von Dimensionshierarchietabellen entwickelt, welche die antizipierbaren² Auswertungsvarianten ermöglichen.“³ „Ziel der Aggregation ist demnach auf Basis konsistenter Dimensionsdefinitionen Summenstrukturen zu erzeugen, die dann aus Gründen der Performanceoptimierung vorberechnet im Data Warehouse

1 [Kem04, S. 31]

2 Im Sinne von vorwegnehmen, also das Zusammenführen von Daten aufgrund einer Dimensionshierarchie

3 [Kem04, S. 31]

gehalten werden können.“¹ Dimensionen wie etwa Kunde oder Produkt werden Hierarchien zugewiesen, die von einer detaillierten Stufe zu einer Gesamtstufe verlaufen, in der alle der Dimension zugehörigen Werte aufsummiert sind. Die Vorberechnung von Kennzahlen kann zum Beispiel als Grundlage dienen um eine Ist-Plan-Abweichung für den Vertrieb zu erfassen. Dazu ist es nötig, Verkaufsdaten die auf der Granularitätsstufe Tag vorliegen auf Jahreswerte zu aggregieren um einen Vergleich mit den Jahresplandaten durchführen zu können.²

Die Dimensionshierarchietabellen werden dabei im Data Warehouse getrennt von den Faktentabellen in Strukturtabellen bzw. Dimension Tables bereitgehalten. Diese unterliegen aufgrund von veränderten Rahmenbedingungen einer dynamischen Entwicklung, so dass die Tabellen modifiziert, gelöscht und neu angelegt werden.³ Als Beispiel kann dafür die Einführung von neuen Produkte bzw. neuer Produkthauptgruppen genannt werden, welche eine Ergänzung der Hierarchietabellen erfordern. Weiterhin ist es bei Aktualisierung der Faktentabellen notwendig, die Daten laut Definition der Dimensionshierarchietabellen zu aggregieren und auf diese Weise performant bereitzuhalten.

Durch die Implementierung von Hierarchien wird das Ziel verfolgt, die Daten anwendungsorientiert bereitzustellen. Das bedeutet, die Daten werden bezüglich zu erwartender Abfragen optimiert und in dieser Form bereitgehalten. „Dies durchbricht die klassische Vorgehensweise der applikationsneutralen Datenmodellierung [...] zugunsten einer stärker auf die antizipierbaren Anwendungen ausgerichteten Datenhaltung.“⁴

Bei der applikationsorientierten Datenhaltung ist bereits die Loslösung vom Muster der strikten Trennung von Daten- und Programmlogik sichtbar. In der vierten Transformationsschicht Anreicherung wird dieser Ansatz weiter verfolgt und so die irreversible Zusammenführung von Daten und Logik vollzogen.⁵ Dabei werden die gefilterten, harmonisierten und aggregierten Daten um funktionale Aspekte erweitert. Funktionale Aspekte stellen in diesem Zusammenhang Berechnungen dar, deren

1 [Cha04, S. 124]

2 Vgl. ebenda, S. 124

3 Vgl. ebenda, S. 124

4 [Cha04, S. 125]

5 Vgl. [Kem04, S. 31]

Ergebnisse die vorhandenen Daten „anreichern“. So werden bereits innerhalb des Transformationsprozesses Kennzahlen ermittelt, die sich auf der Grundlage des vorhandenen Datenmaterial ableiten lassen. Ein weiteres Kriterium für die Berechnung ist die Relevanz der Kennzahl, so dass in erster Linie häufig abgefragte Kennzahlen wie Deckungsbeiträge oder Abweichungen von Planwerten den zusätzlichen Aufwand der Voraggregation rechtfertigen.

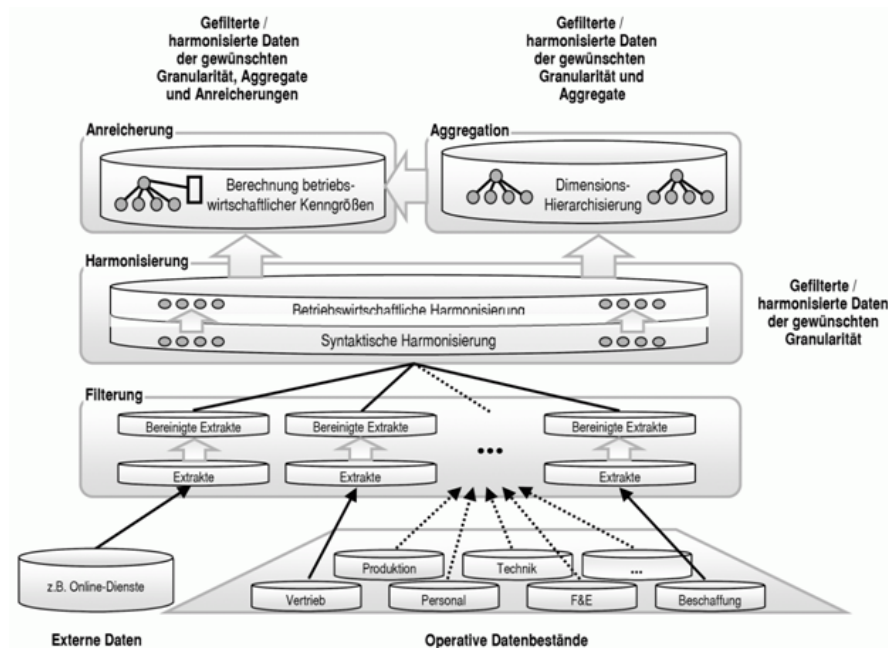


Abbildung 9: Filterung, Harmonisierung, Aggregation und Anreicherung.
[Kem04, S.33]

Grundsätzlich ergeben sich aus dieser Vorgehensweise drei wesentliche Vorteile:

- erhebliche Verkürzung der Antwortzeiten, aufgrund der Voraggregation
- Konsistenz der Daten, aufgrund des einmaligen Berechnungsvorgangs
- Etablierung eines abgestimmten betriebswirtschaftlichen Instruments

Der Transformationsprozess zeigt sich für die Umwandlung von operativen in dispositive Daten als unumgänglich. Die Abbildung 9 zeigt dazu den vollständigen Prozess von der Filterung über die Harmonisierung und Aggregation zur Anreicherung der Daten. Die

Transformationsschichten der Filterung und Harmonisierung gelten dabei als essenziell für die Generierung von dispositiven Daten. Darüber hinaus werden in der Praxis die Schichten Aggregation und Anreicherung hinzugezogen, um eine funktionale Themenausrichtung der Datenbasis zu erhalten.

Im Anschluss an die Einrichtung eines ETL-Prozesses und die initiale Datenübernahme muss unter dem Aspekt der notwendigen Aktualität die Entscheidungsrelevanz der bereitgestellten Daten weiterhin gesichert sein. Zur Aktualisierung der Daten stehen im Wesentlichen drei Konzepte zur Verfügung.¹

Echtzeit-Aktualisierung. Die Echtzeit-Aktualisierung weist eine hohe Komplexität auf, da die Initiative zur Aktualisierung des DWH vom OLTP-System ausgeht, also transaktionssynchron erfolgt. Dazu wird bei Änderung eines Datensatzes im operativen System ein Prozess zur Datenübernahme angestoßen und so die Evidenz der Daten in der Datenbank sichergestellt.

Aktualisierung in Abhängigkeit der Änderungsquantität. Mit diesem Konzept werden die aktualisierten Daten der operativen Systeme erfasst. Bei Erreichung einer definierten Anzahl von Datenänderungen werden diese Daten aufbereitet und in das DWH geschrieben.

Aktualisierung in periodischen Zeitabständen. Die Aktualisierung des DWH in periodischen Abständen kann hingegen zu den bisher aufgeführten Varianten erfolgen, ohne dass eine automatisierte Überwachung der operativen Datenbasis notwendig ist. Ein Update der dispositiven Datenbank erfolgt in periodischen Abständen wie etwa stündlich, täglich oder monatlich, in dem einer auf die Aktualisierung zugeschnittener ETL-Prozess angestoßen wird. Dieser transportiert jedoch Daten, welche bereits im DWH vorliegen und keiner Aktualisierung bedürfen.

Die zeitgesteuerte Evidenthaltung der dispositiven Datenbasis zeigt sich im Vergleich zu den ereignisgesteuerten Varianten *Echtzeit-Aktualisierung* und *Aktualisierung in Abhängigkeit der Änderungsquantität* als wenig komplex und erfordert daher den geringsten Aufwand während der Implementierungsphase. Diese geringe Komplexität begründet den häufigen Einsatz der *Aktualisierung in periodischen Zeitabständen* in

1 Vgl. [Kem04, S. 35]

Unternehmen.

2.1.3 Metadaten eines DWH

Die Informationsobjekte, welche innerhalb des *Data Warehousing* eingelesen, modifiziert und abgelegt werden, sind gekennzeichnet durch Heterogenität und Vielzahl.¹ Die Strukturierung und Beschreibung dieser Objekte erfolgt durch die, den Informationsobjekten zugeordneten Metadaten, welche in einer Repository-Komponente des Data Warehouses abgelegt werden.² Im allgemeinen Verständnis sind Metadaten Daten über Daten, die nach *Bauer* und *Günzel* „beschreibende Informationen über Inhalt, Struktur, Kontext und Bedeutung von Daten, aber auch prozessbezogene Information über die Verarbeitung der Daten“³ enthalten. Ziel ist es, mittels Metadaten Daten besser interpretieren, einordnen, verwalten und nutzen zu können und so etwa über die Herkunft und den Zeitpunkt der letzten Änderung von Daten des Data Warehouses Auskunft zu geben.

Metadaten lassen sich aufgrund ihrer Verwendung einer passiven, aktiven oder semiaktiven Nutzungskategorie zuweisen. Unter die passive Verwendung fallen Metadaten welche Struktur, Entwicklungsprozess und Verwendung von Daten in einem BI-Anwendungssystem konsistent dokumentieren.⁴ Die Dokumentation dient dabei allen Anwendergruppen, Administratoren und Applikationsentwicklern.

Aktive Metadaten werden zur Beschreibung von Prozessen genutzt und können zum Beispiel Transformationsregeln enthalten, die zur Ausführung von Transformations- und Analyseprozessen von einem ETL-Werkzeug interpretiert werden. Semiaktive Metadaten hingegen enthalten Strukturinformationen, die beispielsweise zur Überprüfung von Tabellendefinitionen durch Softwarewerkzeuge herangezogen werden.

Eine weitere Klassifikation der Metadaten berücksichtigt die Zielgruppen. So wird zwischen betriebswirtschaftlich-semanticen und technisch-strukturellen Metadaten unterschieden. Bei technischen Metadaten handelt es sich um eine Datenklasse, die Daten

1 Vgl. [Cha05, S. 33]

2 Vgl. ebenda, S. 33

3 [Bau04, S. 69]

4 Vgl. [Kem04, S. 44]

zur Strukturierung und Beschreibung der Filterungsschicht des ETL-Prozesses bereithält. Bei betriebswirtschaftlichen bzw. Business-Metadaten stehen die Schichten Harmonisierung, Aggregation und Anreicherung sowie die Berechtigungsverwaltung im Vordergrund.¹ Primäres Ziel von Metadaten ist es, das Verständnis und damit die Suche und Nutzung von Daten zu vereinfachen.²

Mit dem Aufbau einer konsistenten Metadatenbasis wird die Minimierung von Aufwand bezüglich des Aufbaus und des Betriebes von BI-Anwendungssystemen verfolgt, da mit Hilfe von Metadaten Prozesse dokumentiert werden können.

Automatisierung der Administrationsprozesse. Durch Scheduling- und Konfigurationsmetadaten kann die Ausführung von BI-Systemprozessen gesteuert und darüber hinaus im Repository Metadaten abgelegt werden, die etwa Fehlerprotokolle, Start- und Abschlusszeit von ETL-Prozessen enthalten.

Flexibler Softwareentwurf. Durch die zentrale Speicherung von Metadaten außerhalb des jeweiligen Werkzeugs im Repository wird die Mehrfachverwendung von Metadaten ermöglicht. Metadaten erleichtern dabei die Arbeit im Umgang mit semantischen und syntaktischen Änderungen sowie den Ausbau eines BI-Anwendungssystems, da sich Modifikationen der Daten direkt auf die beteiligten Prozesse auswirken. Damit kann eine schnelle und widerspruchsfreie Änderung zum Beispiel von Tabellenstrukturen vorgenommen werden, ohne direkt in die Transformationsprozesse einzugreifen.

Schutz- und Sicherheitsaspekte. Weiterhin sollte eine zentrale, prozessübergreifende Verwaltung von Berechtigungsobjekten durch das Repository angestrebt werden, um zu verhindern, dass „jedes einzelne Werkzeug eigene Zugriffsmechanismen verwendet.“³ Auf die Konzeption einer Berechtigungsverwaltung im BI-Umfeld wird im Kapitel 2.5 detaillierter eingegangen.

Metadaten dienen allen Anwendergruppen und sollen einen optimalen Informationsgewinn ermöglichen. In folgenden Punkten verschaffen diese dem Nutzer grundlegende Vorteile:

Datenqualität. Die Qualität von Daten spiegelt sich in Kriterien wie Konsistenz,

1 Vgl. ebenda, S. 44]

2 Vgl. [met10]

3 [Bau04, S. 329]

Korrektheit und Vollständigkeit wider.¹ Mittels Metadaten wird dabei die Bereitstellung von qualitativ hochwertigen Daten durch die Schaffung eines transparenten Prozesses von der Datenquelle bis zur Datenverwendung unterstützt.²

Terminologie und Datenanalyse. Metadaten liefern zu Daten wie etwa Kennzahlen Informationen bzgl. deren Bezeichnung, Herkunft, Abgrenzung und Verwendung. Anwender sollen so in die Lage versetzt werden, präzise Abfragen an das Data Warehouse zu übertragen. Weiterhin ist die zentrale Bereitstellung von Kontextinformationen im Metadaten-Repository Voraussetzung für eine einheitliche Verwendung und Interpretation von Fachbegriffen.

2.2 Datenanalyse

Auf Basis des Datenbestandes einer DWH-Lösung können weiterführende Analysen zur Informationsgewinnung durchgeführt werden. Eine Einteilung der von effizienten und zielorientierten Analysen lässt sich aufgrund der verwendeten Technologien und Konzepte in *Hypothesenverifizierung* und *Hypothesengenerierung* vornehmen.³

Innerhalb der *Hypothesenverifizierung* wird die Gültigkeit einer durch den Anwender getroffene Hypothese untersucht und möglichst belegt. Die dominierende Technologie und das dazugehörige Konzept zur Verifizierung von Hypothesen ist das *On-Line Analytical Processing* und wird in Abschnitt 2.2.1 erläutert.⁴

Mit Hilfe der *Hypothesengenerierung* sollen Beziehungen zwischen Objekten aufgedeckt, dokumentiert und die Ergebnisse dem Benutzer zugänglich gemacht werden. Die Technologie des *Data Mining* ist zur Bewältigung dieser Aufgaben weit verbreitet und wird in Punkt 2.2.2 erörtert.⁵

1 Vgl. ebenda, S. 329

2 Vgl. [Kem04, S. 45]

3 Vgl. [Glu08, S. 143]

4 Vgl. [Möh02, S. 107]

5 Vgl. [Glu08, S. 143]

2.2.1 OLAP

On-Line Analytical Processing ist „eine Software-Technologie, die Managern wie auch qualifizierten Mitarbeitern aus den Fachabteilungen schnelle, interaktive und vielfältige Zugriffe auf relevante und konsistente Informationen ermöglichen soll“¹. In diesem Zusammenhang haben sich dynamische und multidimensionale Sichtweisen auf die Datenbestände etabliert. „Unter Multidimensionalität ist hierbei eine bestimmte Form der Anordnung quantitativer betriebswirtschaftlicher Größen zu verstehen, die relevantes Zahlenmaterial simultan entlang unterschiedlicher Klassen logisch zusammengehöriger Informationsobjekte aufgliedert und dadurch mit der naturgemäß mehrdimensionalen Problemsicht der Unternehmensanalytiker weitgehend korrespondiert“². Die multidimensionalen Sichtweisen verwenden als Datenbasis Daten der operativen Systeme oder als integrierte Business Intelligence Lösung Daten des Data Warehouses.

Codd et al. brachte den Begriff *On-Line Analytical Processing* in die Fachwelt ein. Zuvor wurden bereits innerhalb der Executive Information Systems multidimensionale Sichtweisen verwendet, ohne dass diese jedoch Eingang in die wissenschaftliche Literatur fanden. Mit dem OLAP-Konzept von *Codd* wurde dagegen die Multidimensionalität als zentrales Gestaltungselement in die Diskussion um die IT-basierte Managementunterstützung hineingetragen.³

Codd et al. stellten diesbezüglich zwölf Kriterien auf, welche die OLAP-Fähigkeit eines Informationssystems kennzeichnen. Kritik ernteten die Autoren, da die Kriterien in Zusammenarbeit mit dem Unternehmen *Arbor* erstellt wurden, welches das Datenbanksystem Essbase im Vorfeld der Veröffentlichung auf den Markt brachte. Als negativ wurde dabei in erster Linie die Nichterwähnung dieses Zusammenhangs und die vollständige Erfüllung der Kriterien durch das *Arbor* Produkt betrachtet.⁴

Weitere sechs Regeln, die *Codd* 1995 ergänzte, fanden kaum Beachtung. Trotz der kritischen Auseinandersetzung mit *Codds* zwölf Kriterien, sollen diese hier erwähnt und

1 [Glu08, S. 144]

2 [Cha06, S. 145]

3 Vgl. [Cha06, S. 145]

4 Vgl. [Tot00, S. 61]; [Oeh06, S. 11]

kurz beschrieben werden, nicht zuletzt aufgrund der historischen Bedeutung dieser Regeln, vor allem aber, da diese Kriterien für die Standardisierung von multidimensionalen Systemen verantwortlich sind.

Multidimensionale konzeptionelle Perspektiven. Durch die multidimensionale Sichtweise auf Kennzahlen wird die Fach- oder Führungskraft in die Lage versetzt, intuitiv Analysen entlang der Dimensionen durchzuführen.¹ So kann die Sicht auf Kennzahlen, wie etwa Umsatz oder Lagerbestand, über die Dimensionen Produkt oder Zeit erfolgen.

Transparenz. Der Zugriff auf Daten unterschiedlicher Quellen durch ein OLAP-System muss transparent gestaltet sein. Dem Benutzer wird dabei zur Unterstützung die Herkunft der Daten als Zusatzinformation angeboten. Dennoch darf die Analyse aufgrund der unterschiedlichen Datenquellen nicht negativ beeinflusst werden.²

Zugriffsmöglichkeiten. Ein Informationssystem nach den OLAP-Regeln gewährt den Zugriff auf unternehmensinterne wie auch -externe Datenquellen, unabhängig davon, ob diese relational oder dateibasiert sind. Um die Daten als Basis eines gemeinsamen analytischen Datenmodells zu nutzen, ist die Implementierung von Konvertierungsregeln notwendig.³

Stabile Antwortzeiten bei der Berichterstattung. Die Antwortzeit muss auch bei der Zunahme von Dimensionen oder Datenmengen stabil bleiben.⁴

Client-/ Server-Architektur. Aufgrund der Menge an Daten und der Komplexität der Abfragen ist es notwendig, Daten nicht lokal, sondern zentral zu speichern und über Frontends Zugriff auf diese zu gewähren.⁵

Grundprinzipien der gleichgestellten Dimensionen. „Alle Dimensionen sollen in ihrer Struktur und Funktionalität einheitlich sein.“⁶

Dynamische Handhabung „dünn besetzter“ Matrizen. Dünn besetzte Matrizen „resultieren aus dem Umstand, dass nicht jedes Dimensionselement mit allen Elementen anderer

1 Vgl. [Oeh99, S.29]

2 Vgl. [Glu08, S. 145]

3 Vgl. [Glu08, S. 143]; [Oeh99, S. 30]

4 Vgl. [Gab02, S. 348]

5 Vgl. [Glu08, S. 146]

6 [Bau04, S. 99]

Dimensionen werttragende Verbindungen eingeht.“¹ Um diesem Problem zu begegnen, ist gefordert, dass sich das physische Schema einer mehrdimensionalen Datensicht an die zur Verfügung gestellten Daten anpasst.²

Mehrbenutzerunterstützung. Für OLAP-Systeme ist die Sicherstellung der Mehrbenutzerfähigkeit obligatorisch. Daher muss die Möglichkeit für verschiedene Anwender bestehen, simultan lesend oder auch schreibend auf einen Datensatz zugreifen zu können.³

Unbeschränkte kreuzdimensionale Operationen über Dimensionen hinweg. Innerhalb oder zwischen Würfeln muss eine beliebige Verknüpfung von Daten über Dimensionen hinweg möglich sein. „So können beispielsweise Deckungsbeiträge aus den Umsätzen abzüglich variabler Kosten berechnet werden.“⁴

Intuitive Datenmanipulation. Hinter dem Gedanken der intuitiven Datenmanipulation steht die einfache und ergonomische Gestaltung der Benutzeroberfläche. Zur intuitiven Bedienbarkeit des Werkzeugs stehen Analysemethoden wie Slice and Dice, Roll-Up und Drill-Down zur Verfügung und ermöglichen den „mehrdimensionalen modellierten Datenwürfel entlang der Dimensionen zu analysieren“⁵ Diese Techniken sind Gegenstand dieses Abschnittes.

Flexibles Berichtswesen. Ein OLAP-System ermöglicht die einfache und flexible Erstellung von Berichten. Dazu gehören vordefinierte Standardauswertungen, flexible Adhoc-Berichte und die Möglichkeit auf der Basis von selektierten Daten dynamische Diagramme zu generieren.⁶

Unbegrenzte Dimensions- und Aggregationsstufen. Ein multidimensionales Analysesystem nach Codd verfügt über keine Beschränkungen hinsichtlich der Anzahl von Dimensionen, Relationen und Variablen sowie deren Verdichtung.⁷

Die zwölf Kriterien eines OLAP-Systems nach Codd et al. sind aufgrund der

1 [Cha06, S. 147]

2 Vgl. [Bau04, S. 99]

3 Vgl. [Bau04, S. 99-100]

4 [Oeh99, S. 32]

5 [Ker02, S. 137]

6 Vgl. [Cha06, S. 148]

7 Vgl. ebenda S. 148

herstellerspezifischen Entwicklung kritisiert wurden. Weiterhin ist die nicht realisierte „Trennung zwischen fachlich-konzeptionellen Anforderungen und technischen Realisierungsaspekten“¹ zu bemängeln, da aus den Regeln nicht hervorgeht, welche Speicher- und Datenverwaltungstechniken für die Bereitstellung von mehrdimensionalen Datenmodellen verwendet werden sollen.²

Durch die herstellerabhängige Entwicklung der zwölf OLAP-Regeln war ein Wildwuchs bezüglich deren Ergänzung u. a. durch weitere Hersteller zu verzeichnen, so dass in der Literatur von bis zu 300 Kriterien ausgegangen wird.³ *Pendse* und *Creeth* sahen sich daher veranlasst, herstellerunabhängige Kernelemente einer OLAP-Lösung unter dem Akronym *FASMI* zu veröffentlichen. *FASMI* steht dabei für Fast Analysis of Shared Multidimensional Information.

Fast - Geschwindigkeit. Die Antwortzeit selbst für komplexere Abfragen darf eine Zeitspanne von 20 Sekunden nicht überschreiten. Bei einfachen Abfragen soll die Darstellung der Ergebnisse innerhalb von 1-2 Sekunden erfolgen, so dass sich eine durchschnittliche Wartezeit von ca. fünf Sekunden ergibt.⁴

Analysis - Analyse. Dem Anwender soll es möglich sein, intuitive Analysen auch ohne Programmierkenntnisse durchzuführen. Dazu steht ihm ein breites Spektrum an Auswertungsmöglichkeiten sowie erweiterte Navigationshilfen wie *Drill-Down* und *Roll-Up* zur Verfügung.⁵

Shared – Mehrbenutzerfähigkeit. Wie in den OLAP-Regeln fordert auch *FASMI* die Möglichkeit des Mehrbenutzerbetriebes. Dafür ist die Implementierung von Sicherheitsmechanismen und einer effektiven Zugangssteuerung notwendig.⁶

Multidimensional. *FASMI* stellt die für OLAP-Systeme essentielle Forderung nach einer mehrdimensionalen Sicht auf die Informationsobjekte.

Information – Datenumfang. *Chamoni* und *Gluchowski* konstatieren dazu, dass „ein

1 [Cha06, S. 148]

2 Vgl. [Cha06, S. 148]

3 Vgl. [Kem04, S. 94]

4 Vgl. [Hol99, S. 54]

5 Vgl. [Knö06, S. 58]

6 Vgl. [Kem04, S. 94]

System [...] dann von hohem Nutzwert [ist], wenn es mehr Datenelemente bei stabiler Antwortzeit analysieren kann.“¹ Der Erfolg eines OLAP-Systems ist also abhängig von der Verfügbarkeit der Daten und Ressourcen.

Bei der Analyse von Unternehmenskennzahlen im mehrdimensionalen Raum werden mehrere Kriterien herangezogen. Daher können Fragestellungen beantwortet werden wie etwa: In welchem Zeitraum wurde mit welchem Artikel in welcher Region wie viel Umsatz erzielt? In diesem Beispiel dienen das Gebiet und der Zeitraum als Kriterium bzw. Dimension und der Umsatz als Kennzahl. Zur Verdeutlichung werden diese Analysen in der Literatur in Form eines dreidimensionalen Datenwürfels dargestellt. In der Praxis befinden sich jedoch multidimensionale Datenmodelle mit 15 bis 20 Kriterien im Einsatz und *Codd* fordert sogar eine unbegrenzte Anzahl von Dimensionen.² Daher könnte das Beispiel um die Dimension des Mitarbeiters erweitert werden. Die Frage würde dann lauten: Welcher Mitarbeiter hat in welchem Zeitraum mit welchem Artikel in welcher Region wie viel Umsatz erzielt? Die Kennzahlen wie die Umsatzzahlen des Beispiels liegen dabei in atomarer Form vor und werden entlang der Dimensionen verdichtet.

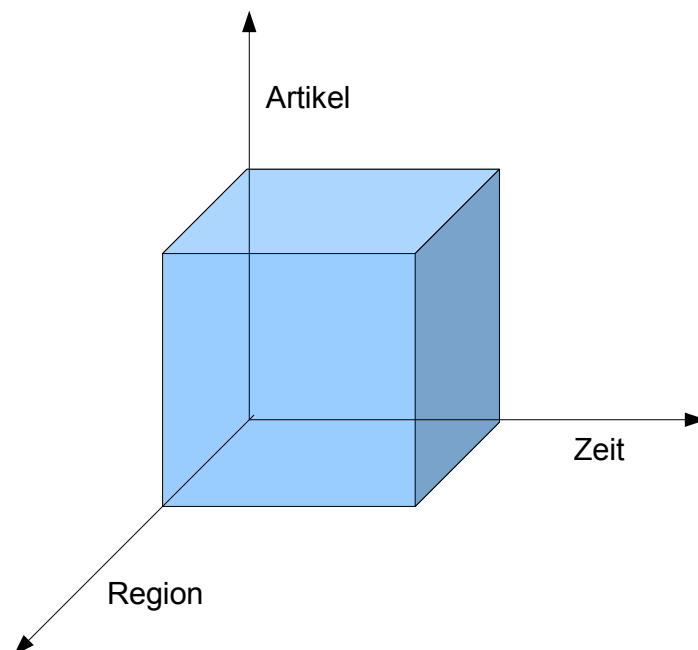


Abbildung 10: Dreidimensionales Datenmodell

¹ [Cha06, S. 150]

² Vgl. [Cha06, S. 148]

Um Analysen auf multidimensionalen Modellen durchführen zu können, war es erforderlich, bisherige Herangehensweisen aufzugeben. So wurden neue Techniken benötigt, um „innerhalb der durch das Datenmodell aufgespannten multidimensionalen Würfel zu navigieren und komplexe Operationen über mehrere Dimensionen hinweg durchzuführen.“¹

Drill-Down. Das *Drill-Down* bezeichnet eine vertikal abwärts gerichtete Navigationstechnik im mehrdimensionalen Raum, bei welcher die Kennzahlen entlang der Dimensionen bis zum atomaren Zustand aufgebrochen werden. Besonders hilfreich ist diese Methode für die Ursachenforschung, so dass etwa beim sprunghaften Anstieg der Umsatzzahlen die verantwortlichen Artikel ausfindig gemacht werden können.

Roll-Up bzw. Drill-Up. Das Gegenteil bewirkt die Verwendung der Navigationstechnik *Roll-Up* bzw. *Drill-Up*. Dabei werden Daten entlang der Dimensionen aggregiert, um so etwa die Frage zu beantworten, welche Auswirkung die Umsatzsteigerung des Vertriebsbüros im Bundesland Sachsen auf die Umsätze in Deutschland hat.

Rotate. Durch das *Rotieren* bzw. *Drehen* des Würfels ist es möglich, eine Kennzahl aus unterschiedlichen Perspektiven zu betrachten. So kann im verwendeten Beispiel bei Analyse der Umsatzzahlen in einem Zeitraum ein Perspektivwechsel von Artikel nach Region erfolgen.

Slice. Mittels der Analysetechnik *Slice* wird dem Würfel eine Schicht entnommen, indem ein Filter auf einen Attributwert einer Dimension gesetzt wird. Auf diese Weise kann zum Beispiel der Umsatz eines bestimmten Artikels untersucht werden.

¹ [Knö06, S. 59]

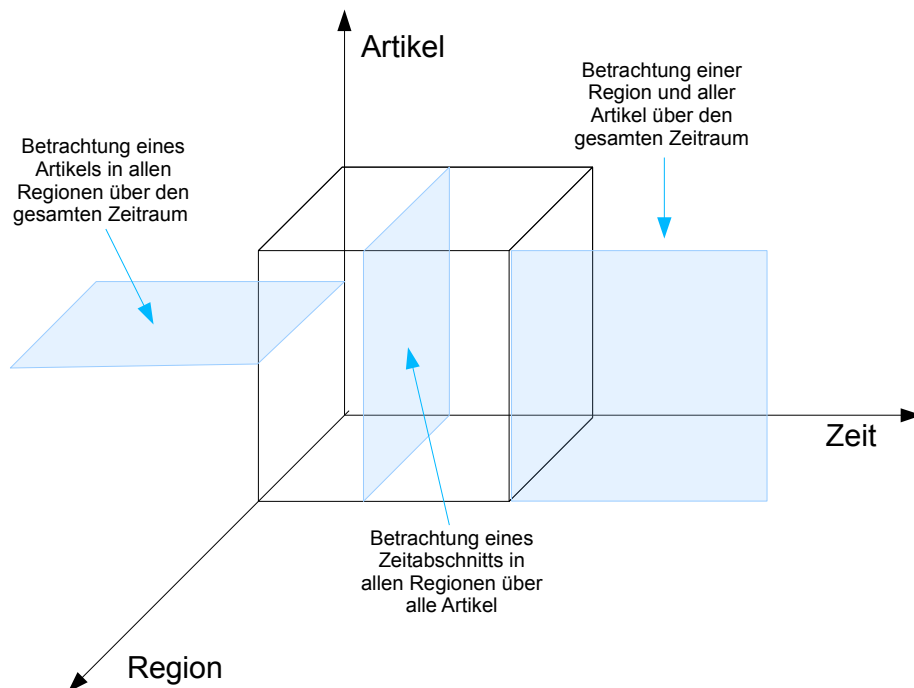


Abbildung 11: Analysetechnik Slice

Dice. Durch die Einschränkung von n Dimensionen auf bestimmte Ausprägungen der Attribute, wie Artikel, Region und Zeit, wird ein mehrdimensionales Datenmodell generiert, welches Bestandteil des Hypercubes¹ ist. Auf der Basis dieses Würfels ist es möglich, weitere Analysen durchzuführen.

Split and Merge. Die Methode *Split* wird herangezogen, um Dimensionen eines Datenwürfels in eine Auswertung einzubeziehen. Mit der Technik *Merge* kann dieser Vorgang rückgängig gemacht werden.

Drill-Through & Drill-Across. „*Drill-Through* and *Drill-Across* sind besondere Operationen, da sie eine Navigation über den originären multidimensionalen Datenraum hinaus ermöglichen.“² Die Analysetechnik *Drill-Through* setzt an den Grenzen von *Drill-Down* an. Erreichen die analysierten Daten die Stufe der geringsten Granularität, erfolgt der automatische Wechsel zur operativen Datenbasis. Dieser Übergang muss ohne Beeinflussung der Analyse durchführbar sein.

¹ Als Hypercube wird in diesem Zusammenhang das ungefilterte multidimensionale Datenmodell bezeichnet

² [Kem04, S. 99]

Die Methode *Drill-Across* benötigt und verwendet Schnittmengen zwischen OLAP-Würfeln, über die der horizontale Wechsel zwischen verschiedenen mehrdimensionalen Datenmodellen erfolgen kann. Aufgrund dieser erforderlichen Schnittmengen ist der Übergang zu einem anderen Würfel daher nur möglich, wenn die dimensionalen Hierarchien wiederholt auftreten.¹

2.2.2 Data Mining

Wörtlich übersetzt bedeutet *Data Mining* „Schürfen oder Graben“ in Daten. Während mit Hilfe der beschriebenen OLAP-Systeme Hypothesen überprüft werden können, sollen durch das Data Mining automatisch Hypothesen generiert werden. Es wird hierbei davon ausgegangen, dass sich in Unternehmen vielfältige und große Datenmengen angesammelt haben, hinter denen sich aufschlussreiche Informationen verbergen können, wenn die Daten ausgewertet und zum Vorteil des Unternehmens eingesetzt werden.²

Zu den häufigsten Anwendungsgebieten des Data Mining gehören nach *Chamoni et al.* die Klassifikation, das Clustern und das Aufdecken von Abhängigkeiten.³ Um ein Objekt zu klassifizieren, wird es einer oder mehreren vorher festgelegten Kategorien zugeordnet. Beim Clustering werden Daten zwar auch in Klassen bzw. Cluster geordnet, jedoch liegen hierbei im Gegensatz zur Klassifikation im Vornherein keine festgelegten Kategorien vor. Aufgabe der Data Mining-Anwendungen ist die Zuordnung der Daten zu Clustern. Um Korrelationen zwischen einzelnen Informationsobjekten zu bewerten, müssen Abhängigkeiten aufgedeckt werden. Bei interdependenten Objekten wird anschließend ein Korrelationskoeffizient gebildet, um die Stärke der Abhängigkeit quantifizieren zu können.

1 Vgl. [Sch07, S. 121]

2 Vgl. [Cha05, S. 22]

3 Vgl. [Cha05, S. 22]

2.3 Portale

BI-Portale werden verwendet, um die Ergebnisse der Phasen *Bereitstellung* und *Entdeckung* dem adressierten Nutzerkreis unter einer einheitlichen Oberfläche zur Verfügung zu stellen. Der Terminus *Portal* bezieht sich dabei auf den lateinischen Begriff *Porta* und hat seinen Ursprung im *World Wide Web*. In diesem Zusammenhang ist ein *Webportal* eine Webseite und zentraler Einstiegspunkt, um verschiedene Dokumente aus dem Internet abzurufen. Neben Webportalen gibt es verschiedene Ausprägungen von Portalen wie Enterprise Information Portale, welche eine übergeordnete Kategorie der hier zu thematisierenden *Business Intelligence Portale* darstellen.

BI-Portale bieten einen definierten Einstiegspunkt, welcher Nutzern einen integrativen Zugang zu Informationen ausgewählter Themengebiete sowie Anwendungen gewährt, auf deren Basis fundierte Entscheidungen getroffen werden können.¹ Die Zugriffsmöglichkeit auf dieses breite Angebotsspektrums soll dabei durch eine geringe Komplexität gekennzeichnet sein, so dass *Business Intelligence Portale Single-Sign-On* unterstützen. Dadurch erhält der Benutzer nach einmaliger Authentifizierung Zugangsberechtigung zu den für ihn relevanten Inhalten und Anwendungen des Portals. Dem Nutzer wird dabei eine Rolle zugeordnet, welche die Personalisierung der Applikation etwa durch spezifisches Design aber auch durch Restriktionen der Berechtigungen ermöglicht.

Mit Hilfe von *BI-Portalen* können Benutzer auf Informationen zugreifen, die durch Standardberichte, Cockpits bzw. Dashboards visualisiert und tabellarisch oder grafisch aufbereitet werden.² Während Standardberichte primär Informationen zur Durchführung und Lenkung des operativen Geschäfts enthalten, umfassen Cockpits bzw. Dashboards komprimierte und wesentliche Informationen zur Unterstützung des Topmanagements. Cockpits bzw. Dashboards können weiterhin als Ausgangspunkt für tiefgreifende Untersuchungen der angezeigten Informationen etwa durch den gezielten Zugriff auf BI-Anwendungen wie *OLAP* oder *Data Mining* genutzt werden.

Portale sind für die direkte Nutzung am Bildschirm auf der Basis von *Thin-Clients*

¹ Vgl. [Glu08, S. 216-217]

² Vgl. [Cha06, S. 101]

optimiert. Diese beschränken sich dabei auf die Darstellung von angeforderten Informationen innerhalb von Webbrowsern, so dass die zusätzliche Installation von Client-Software nicht notwendig ist. Grundlage für die Präsentation der Inhalte durch diese Clients ist dabei die Trennung der Präsentationsschicht von der Logik- und Datenschicht, so dass Aufgaben zur Speicherung und Verarbeitung von Servern übernommen werden und damit die Darstellung der Inhalte in Webbrowsern von Netzwerkcomputern erfolgen kann. Zusätzlich ist es heute aufgrund der dynamischen Wirtschaftsentwicklung erforderlich, dass Führungskräfte kontinuierlich mit Informationen des internen und externen Unternehmensumfelds versorgt werden und damit gebündelte und entscheidungsrelevante Inhalte auch innerhalb von mobilen Endgeräten wie Smartphones präsentiert bekommen.

2.4 Standardsoftware und BI-Markt

Um den Aufwand bei der Einführung eines betriebswirtschaftlichen Anwendungssystems zu minimieren, greifen Unternehmen zunehmend auf Standardsoftware zurück¹ und entscheiden sich somit gegen die Implementierung von Individualsoftware, welche speziell für das Unternehmen entwickelt und so auf dessen Bedürfnisse zugeschnitten ist. Die Erstellung von Individualsoftware birgt hohe Risiken bezüglich der anfallenden Kosten sowie der Einhaltung des definierten Zeitrahmens, so dass Unternehmen aufgrund dieser Unsicherheiten vornehmlich auf Standardsoftware setzen.

„Allgemein gilt eine Softwarelösung als Standardsoftware, wenn

- sie in Unternehmen verschiedener Struktur eingesetzt werden kann
- sie an diese Unternehmen in gewissem Maße angepasst werden kann, ohne den Programmcode zu ändern (Customizing)
- die Entwicklung der Software abgeschlossen ist und der Beschaffungspreis feststeht.“²

Auch im Umfeld von *Business Intelligence* bedienen sich Unternehmen der

¹ Vgl. [Glu08, S. 10]

² [Dis03, S. 385]

Standardsoftware und finden in diesem Segment ein breites Anwendungsspektrum vor, welche alle Bereiche des engeren, erweiterten und analyseorientierten BI-Verständnis abdecken. Innerhalb dieses Marktes bewegen sich Anbieter, welche Unternehmen mit speziellen BI-Lösungen unterstützen und Anbieter die essentielle BI-Anwendungen unter einer Plattform bzw. Suiten vereinen.

Spezialanbieter konzentrieren sich auf separate Aufgaben der BI wie Datenintegration oder Data Mining und zeichnen sich innerhalb dieser funktionalen Schwerpunkten mit hoher Kompetenz aus. Im Segment der Datenintegration sind dabei etwa die Produkte *ETL-Manager* von *Iway* oder *CO>Operating System* von *Ab Initio* zu erwähnen. Eine weitere Ausprägung dieser Kategorie und Gegenstand der Diplomarbeit ist die Open Source Software *Talend Open Studio*.¹ Entwickelt wird diese Applikation durch das Unternehmen *Talend* mit Sitz in Suresnes in Frankreich und Los Altos in den USA, welches sich als Anbieter kommerzieller *Open Source Software* zur Datenintegration etabliert hat.

Der Begriff Open Source Software kann mit freier Software übersetzt werden. Dies bedeutet nicht zwangsläufig, dass diese kostenlos zur Verfügung steht, vielmehr beschreibt die Bezeichnung, dass der Quelltext der freien Software öffentlich zugänglich ist, so dass „jeder das Programm im Ursprung lesen, analysieren, verstehen und bei Erfordernis seinen eigenen Wünschen anpassen“² kann.

Mit dem Werkzeug *RapidMiner* bietet das Unternehmen *Rapid-I GmbH* eine weitere freie Software, mit deren Funktionsumfang Hypothesenverifizierung, also Data Mining, betrieben werden kann. *RapidMiner* ist dabei das am häufigsten genutzte Open-Source-Tool für Data Mining und wurde vom CEO des Unternehmens im Rahmen einer Dissertation erstellt. Die thematisierte Software wird in der Community-Version über das Webportal *SourceForge* zum kostenfreien Download angeboten. Um laufende Projekte zu finanzieren, bietet die *Rapid-I GmbH* darüber hinaus kostenpflichtigen Support und Versionen³ der Software, welche zusätzliche Komponenten etwa zur Darstellung der Ergebnisse innerhalb von Webbrowsern enthalten. Neben den Open-Source-Produkten zum Data Mining, wie *RapidMiner* und *KNIME*, existiert ein breites Angebot an

1 Internetauftritt: <http://de.talend.com>

2 [Sal05, S. 6]

3 Siehe <http://rapid-i.com/content/view/182/190/>

proprietären Produkten wie etwa der *Enterprise Miner* von *SAS*, *Statistica* von *StatSoft* oder *NetMind* von *MindLab*.

Zusätzlich zu diesen speziellen Applikationen des BI werden sogenannte Business Intelligence Plattformen bzw. BI-Suiten angeboten. Diese erheben den Anspruch, ein breites Spektrum von möglichen Anwendungen im BI-Umfeld abzudecken, so dass Unternehmen nur bedingt vor die Herausforderung gestellt werden, aus verschiedenen Softwareangeboten eine eigene BI-Suite zu entwickeln oder ggf. Referenzmodelle einzukaufen. Dennoch sieht sich der potentielle Kunde mit einer Vielzahl von Anbietern konfrontiert, welche mitunter verschiedene BI-Anwendungen unter einer Plattform vereinen. Um eine Vergleichbarkeit der Angebote zu ermöglichen, definierte das Unternehmen *Gartner Incorporated*, ein Anbieter von IT-Marktforschungsergebnissen, Funktionen, die eine Business Intelligence Plattform bereitstellen muss:

Integration. Die Plattform bietet verschiedene Anwendungen und integriert diese über eine zentrale Metadatenverwaltung und Administration.

Informationsversorgung. Zur Managementunterstützung sind Informationen über Reporting-Werkzeuge, Dashboards und Ad-hoc-Queries abrufbar. Weiterhin wird die Integration von Microsoft Office Produkten gefordert sowie die Bereitstellung einer Suchfunktion.

Analyse. Funktionen zur Analyse bieten insbesondere Scorecards, OLAP und Data Mining.

Die Kriterien sind hier zusammengefasst dargestellt. In Anlage B werden die Veröffentlichung der Gartner Inc. vom Januar 2010, in welcher die Forderungen an eine Business Intelligence Plattform definiert und detailliert behandelt werden. Auf Basis dieser Forderungen wurde, ebenfalls durch die Gartner Inc., das folgende Portfolio erstellt, in dem alle Softwareanbieter erscheinen, welche die genannten Forderungen erfüllen. Die Anbieter sind in die vier Quadranten eines Koordinatensystems geordnet, welche folgende Bezeichnungen tragen: challengers (Herausforderer), leaders (Marktführer), niche players (Nischenspieler) und visionaries (Visionäre).

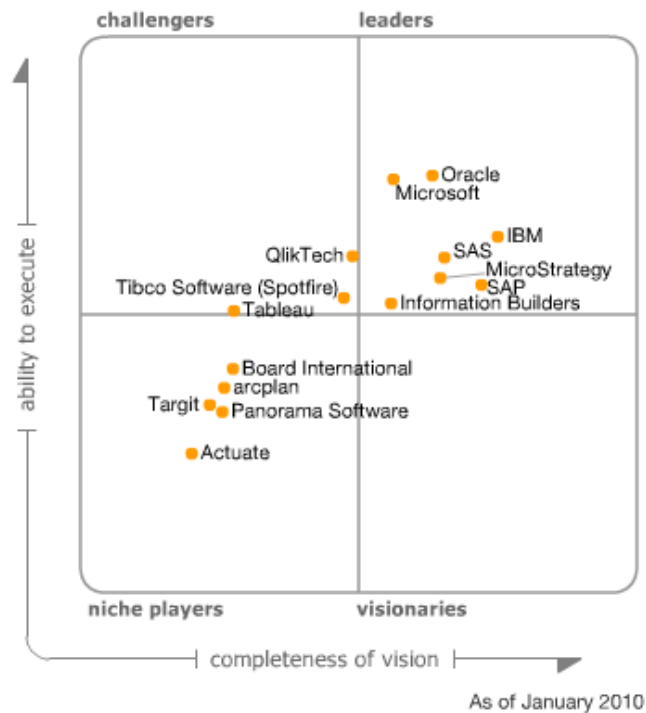


Abbildung 12: BI-Anbieterportfolio. [Sal10]

Der BI-Markt für Standardsoftware geriet um das Jahr 2006 durch verschiedene Zukäufe von IT-Konzernen in Bewegung, so dass eine Marktkonsolidierung zu verzeichnen ist. Aufgrund der kontinuierlichen Wachstumsraten in diesem Segment kaufte etwa *SAP* mit *Business Objects* einen der führenden Anbieter von BI-Lösungen und gab dafür die bisherige Unternehmensphilosophie Wachstum aus eigener Kraft zu generieren auf. Weiterhin steht *ORACLE* vor der Aufgabe die Anwendungen von *Hyperion* ins eigene Produktportfolio zu integrieren. *IBM* hingegen verschaffte sich durch den Zukauf des kanadischen Unternehmens *Cognos* Zugang zu Know-How und Märkten.¹ *Cognos* machte zuvor im Oktober 2007 mit dem Erwerb von *Applix* auf sich aufmerksam.

¹ Vgl. [And10]

Rang	Anbieter	BI-Softwareumsatz in Deutschland 2008 in Mio. Euro	Veränderung zum Vorjahr	Marktanteil
1	SAP	110	8,90%	14,60%
2	ORACLE	104	7,80%	13,80%
3	SAS	89	4,70%	11,80%
4	IBM	72	-1,40%	9,50%
5	Microsoft	53	20,00%	7,10%
	TOP 5	428	8,00%	56,80%
	Gesamt	754,1	6,20%	100,00%

Tabelle 2: Verteilung des BI-Marktes in Deutschland. Nach [Ale10]

Der deutsche Markt wird durch die Angebote von *SAP*, *ORACLE*, *SAS*, *IBM* und *Microsoft* dominiert, welche laut einer Studie des *Business Application Research Center* in Summe etwa 60 Prozent Marktanteil erreichen.

Mit BI-Standardsoftware sollen zahlreiche Aufgaben der Datensammlung, -aufbereitung und Verteilung automatisiert werden und damit Informationen den Führungskräften aller Ebenen zur Entscheidungsunterstützung zugänglich gemacht werden. Die Aufgabe von Standardsoftware dabei ist, Systeme bereitzustellen, welche Implementierung und Betrieb eines BI-Systems mit geringem Aufwand ermöglichen. Dazu finden Unternehmen ein breites Softwareangebot vor, das neben speziellen Nischenlösungen auch umfassende Lösungen in Form von Business Intelligence Plattformen bereithält.

2.5 Einführung einer BI-Lösung mit Sicherheitskonzept

Bei der Einführung einer Business Intelligence Lösung wird das Projektteam nach *Suhr/Suhr* mit folgenden drei Problemfeldern konfrontiert und steht vor der Herausforderung, diese zu bewältigen.¹

- „Beherrschung der Komplexität von Aufgabenstellung und Problemlösung
- Bewältigung der Dynamik von Anforderungen und Einsatzbedingungen

¹ Vgl. [Suh93, S. 96]

- Beteiligung betroffener Interessengruppen.²

Zur Lösung dieser Schwierigkeiten wurden im Bereich des Software-Engineering Phasen- und Vorgehensmodelle wie das *phasenorientierte BI-Vorgehensmodell*, das *parallel iterative BI-Phasenmodell* und das *erweiterte BI-Vorgehensmodell* entwickelt. Während das *phasenorientierte BI-Vorgehensmodell* die Entwicklungsschritte Analyse, Design und Produktivsetzung enthält, weist das *parallel iterative BI-Phasenmodell* darüber hinaus vor- und nachgelagerte Aktivitäten auf. Das *erweiterte BI-Vorgehensmodell* bezieht überdies prozessbegleitende Elemente wie die BI-Strategie, das BI-Qualitätsmanagement und das BI-Projektmanagement mit ein. Deshalb wird auf Basis dieses zuletzt genannten Modells das Vorgehen zum Aufbau und Einsatz einer BI-Lösung im Unternehmen aufgezeigt.

Die BI-Strategie stellt in diesem Zusammenhang eine übergeordnete Komponente des BI-Vorgehensmodells dar, welche sich aus der Unternehmensstrategie ableitet. Die Aufgaben zur Koordinierung und Planung der einzelnen Aktivitäten obliegen dem Projektmanagement, während die Ergebnisse dieser Aktivitäten durch das Qualitätsmanagement verifiziert werden.

Der gesamte Lebenszyklus bezüglich der Gestaltung und dem Betrieb einer BI-Lösung sowie die notwendigen Aktivitäten sind in Abbildung 13 dargestellt und werden von der BI-Projektdefinition bis zum BI-Betrieb und -Wartung konkretisiert.

2 [Glu08, S. 254]

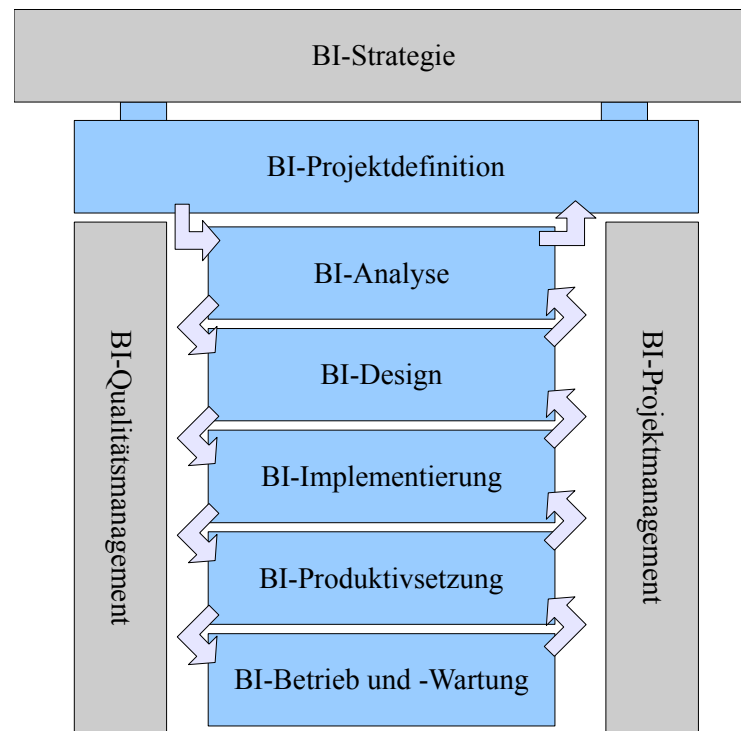


Abbildung 13: Schema des erweiterten BI-Vorgehensmodells.
[Gab09, S. 66]

Projektdefinition. Nachdem die strategische Entscheidung zur Einführung eines BI-Systems getroffen wurde, werden innerhalb der Projektdefinition die Projektziele und die Projektorganisation festgelegt sowie ein erster Projektplan erstellt.

Analyse. Die Analysephase ist gekennzeichnet durch eine Vielfalt an Einzelaktivitäten wie die Ist-Analyse, Schwachstellenanalyse, Bedarfsanalyse, Quelldatenanalyse und semantische Modellierung. Die Resultate dieser Einzelaktivitäten wie die funktionalen und inhaltlichen Anforderungen an das System sowie die Quellen der benötigten Daten und die Verknüpfung dieser Daten im Geschäftsprozess werden in einem Dokument wie einem Lastenheft festgehalten, das die Basis für die weitere Vorgehensweise bildet. Das Lastenheft wird dabei vom potentiellen Auftraggeber erstellt und enthält unter anderem anzustrebende Ergebnisse sowie ausformulierte Anforderungen an Soft- und Hardware, ohne ein Produkt zu präferieren.

Design. Die Phase des BI-Designs umfasst Aktivitäten wie die Auswahl und Festlegung von Systemarchitektur sowie aller Hard- und Softwarekomponenten. Ebenso lassen sich

der Designphase sowohl die Gestaltung des Webportals und des ETL-Prozesses als auch die Erstellung eines logischen Datenmodells zuordnen.

Implementierung. Grundlage für die BI-Implementierung sind die Ergebnisse der Analyse- und Designphase. In diesem Zusammenhang werden die Resultate bis zur lauffähigen BI-Lösung innerhalb der ausgewählten Hard- und Softwarekomponenten in einer Test- und Entwicklungsumgebung umgesetzt. Wesentliche Aufgabe der BI-Implementierung stellt der Entwurf eines Sicherheitskonzeptes dar, welches im weiteren Verlauf des Kapitels präzisiert wird.

Produktivsetzung. Im Rahmen dieser Phase werden die Ergebnisse der Test- und Entwicklungsumgebung in eine Produktivumgebung überführt und damit für die Endanwender freigegeben.

Betrieb und Wartung. Die Phase des BI-Betriebes und der BI-Wartung erstreckt sich von der Produktivsetzung bis zur Inbetriebnahme des Systems. Bei erfolgreicher Einführung einer Business Intelligence-Software sehen sich die mit administrativen Aufgaben betrauten Mitarbeiter mit einer Vielzahl an Anforderungen von Fachabteilungen konfrontiert. Ziel ist die ständige Verbesserung der Systemleistungsfähigkeit, der Funktionalität und der Inhalte.

Nunmehr wird das oben erwähnte Sicherheitskonzept bzw. die Implementierung von Berechtigungsstrukturen konkretisiert. Die gewissenhafte Konzeptionierung ist gerade auch im Umfeld von Business Intelligence bedeutsam, da das System unternehmens- und personenkritische Daten wie Umsatzzahlen, Prognosen und Arbeitsentgelte beinhaltet. Um Aufgaben der Berechtigungsadministration mit möglichst geringem Aufwand und dennoch wirksam bewältigen zu können, ist es von Vorteil, wenn BI-Konzepte mit integrierter Datenhaltung eine zentrale Vergabe von Zugriffsrechten für alle Anwendungen der BI-Lösung erlauben. Dazu greifen viele Systeme auf eine rollenbasierte Berechtigungsadministration zurück. Hierbei werden Benutzern bzw. Benutzergruppen Rollen zugewiesen. Diese Zuweisung leitet sich aus den Aufgaben und Verantwortungsbereichen der Anwender ab. Eine Benutzerrolle umfasst in diesem Zusammenhang alle Berechtigungen, die der Anwender benötigt, um die ihm zugewiesenen Aufgaben im Unternehmen zu erfüllen. Diesbezüglich wird innerhalb von

BI-Systemen zwischen Zugriffsrechten auf spezifische Daten und Zugriffsrechten für spezifische Anwendungen differenziert.

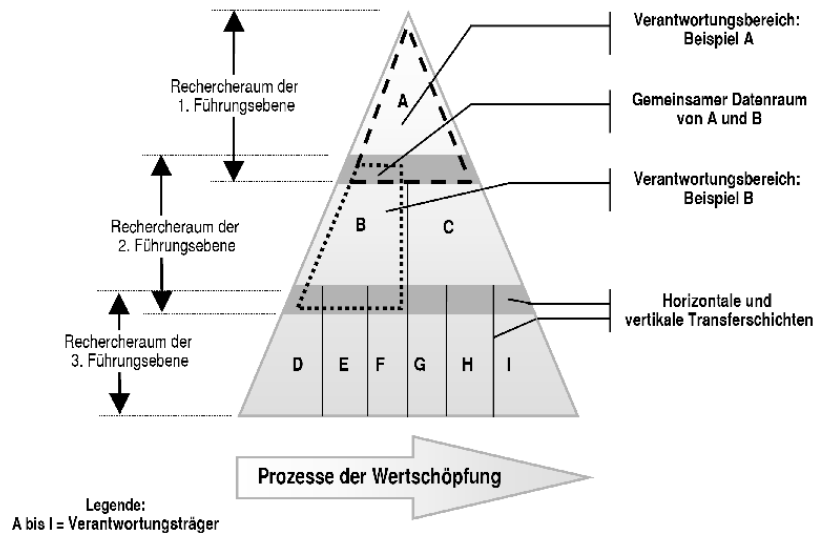


Abbildung 14: Modell zur Berechtigungskonzeption. [Kem04, S. 51]

Abbildung 14 zeigt ein Modell zur Berechtigungskonzeption eines BI-Systems unter Einbeziehung der ersten, zweiten und dritten Führungsebene eines Unternehmens. Durch die Zuweisung von spezifischen Datensichten kann nach *Kemper et al.* sichergestellt werden, „dass alle Anspruchsgruppen durch eine Beschränkung des dispositiven Datenzugriffs auf bestimmte Tabellenfelder von Basis- oder Zusammenfassungstabellen nur mit ihrem aufgabenspezifischen Datenmaterial arbeiten können.“¹ Auf diese Weise wird gewährleistet, dass die Datensicht für Anwender in der horizontalen Recherchebreite wie auch in der vertikalen Recherchetiefe begrenzt bleibt. So wird zum Beispiel einer Führungskraft der dritten Ebene im Aufgabenbereich *Produktion* kein Einblick in Personaldaten gewährt. Ebenso werden Zugriffsrechte für das Topmanagement innerhalb der vertikalen Recherchetiefe limitiert, so dass dieses lediglich Zugang zu verdichteten Personaldaten erhalten.

1 [Kem04, S. 50]

3 Konzept und Implementierung einer BI-Lösung bei AVANCIS

Gegenstand der Diplomarbeit ist die Einführung einer *Business Intelligence Lösung* bei der AVANCIS GmbH & Co. KG am Standort in Torgau, Sachsen, wobei als Grundlage zur Beschreibung und Analyse des Einführungsprozesses das erweiterte *BI-Vorgehensmodell* dient.

Die Phasen des erweiterten BI-Vorgehensmodells, *Projektdefinition*, *Analyse* und *Design*, sind der Implementierungsphase der Einführung einer BI-Lösung vorgelagert. An die Implementierung schließen sich die *Produktivsetzung* sowie *Betrieb* und *Wartung* an. Die Beschreibung und Analyse dieser Phasen erfolgt unter Verwendung und Einbezug der im Unternehmen ausgewählten Software *STAS CONTROL*.

Die *STAS GmbH* bietet Möglichkeiten, kaufmännische Daten in vordefinierten Analysen und Auswertungen bereitzustellen und darüber hinaus zusätzliche Standardberichte, vor Ort beim Auftraggeber individuell auf Basis des Lastenheftes zu erstellen. Aufgrund dieser Zusagen wurde die *STAS GmbH* beauftragt ihre Business Intelligence Lösung *STAS CONTROL* bei AVANCIS zu implementieren.

Das Unternehmen STAS aus Reilingen hat sich mit dem Softwarekonzept *STAS CONTROL* das Ziel gesetzt, Business Intelligence Lösungen innerhalb eines kurzen Projektintervalls von 10 Tagen aufwärts bei ihren Kunden umzusetzen. *STAS CONTROL* bedient sich dabei der modifizierten Talendversion *STAS CONTROL ETL powered by talend¹* und *IBM Cognos 8.4*.

Um die Einführung einer BI-Lösung innerhalb kürzester Zeit zu realisieren, bietet die *STAS GmbH* „schlüsselfertige betriebswirtschaftliche Softwarelösungen zur Planung, Analyse und Steuerung mittelständischer Unternehmen“² an.

Die Konstante im System *STAS CONTROL* stellt dabei der Bedarf an betriebswirtschaftlichen Kennzahlen und deren Auswertungsmöglichkeiten dar, so dass

¹ Im weiteren Verlauf der Arbeit STAS CONTROL ETL genannt.

² [STA10]

STAS CONTROL vordefinierte Berichte, Cockpits, Dashboards und Analysen in *Cognos 8.4* zu betriebswirtschaftlichen Teilgebieten wie Finanzwirtschaft, Controlling, Fertigung und Materialwirtschaft sowie abteilungsübergreifende Auswertungen, die für das Topmanagement konzipiert sind, enthält. Diese in *Cognos 8.4* vorkonfigurierten Auswertungen nutzen als Datenquelle ein Data Warehouse oder auf dem Datenbestand eines Data Warehouses erzeugte OLAP-Würfel. Der Leistungsumfang von *STAS CONTROL* bietet daher, auf die vordefinierten Auswertungen zugeschnittene, Metadaten für die strukturelle Einrichtung eines Data Warehouses und mittels *IBM Cognos Transformer* vorkonfigurierte und nach Anwendungsgebiet separierte OLAP-Würfel.

Die Schnittstelle zwischen den dispositiven Datenquellen und den über *Cognos Connection* zugänglichen Auswertungen bildet der *Content Store*, welcher mit Metadaten über den *Frameworkmanager* gefüllt wird. Die unter dem Begriff *Packages* organisierten Metadaten liegen innerhalb von *STAS CONTROL* vorkonfiguriert im *Frameworkmanager* und schließlich in publizierter Form im *Content Store* vor.

Die beim Kunden vorhandenen und damit abzubildenden Quellsysteme können hingegen als Variable im System *STAS CONTROL* angesehen werden. Dabei sieht sich die *STAS GmbH* gerade bei mittelständigen Auftragsunternehmen einem breiten Spektrum an operativen Systemen ausgesetzt, welche wiederum durch zahlreiche Customizing-Einstellungen individualisiert und so auf die Bedürfnisse des Kunden zugeschnitten sind.

Mit den definierten Variablen und Konstanten stehen ein vordefiniertes dispositives System und ein oder mehrere operative Systeme gegenüber, so dass sich aus dieser Konstellation die Aufgabe ableitet, eine Schnittstelle zwischen diesen Systemen zu konfigurieren und damit einen ETL-Prozess zur Befüllung des Data Warehouses zu implementieren. Zur Realisierung dieser Datenschnittstelle wird innerhalb von *STAS CONTROL* das Werkzeug *STAS CONTROL ETL powered by Talend* eingesetzt. Mit den Teilprozessen Filterung, Harmonisierung, Aggregation und Anreicherung findet eine auf die Struktur des Data Warehouses zugeschnittene Transformation von operativen Daten in dispositive Daten statt.

STAS CONTROL stellt demzufolge ein Softwarekonzept dar, das sich im Wesentlichen aus den Komponenten *Cognos 8.4*, *STAS CONTROL ETL* und den Metadaten für ein Data

Warehouse zusammensetzt. In *Cognos 8.4* liegen zusätzlich von der *STAS GmbH* vordefinierte Auswertungen in Form von Ad-hoc-Analysen, Ad-hoc-Abfragen, Standardberichte, Dashboards und Cockpits vor, die wiederum auf vorkonfigurierten relationalen oder multidimensionalen Datenmodellen basieren, so dass beim Kunden, neben der Softwareinstallation und -inbetriebnahme, lediglich die Ausgestaltung des ETL-Prozesses vorgenommen werden muss.

3.1 Projektdefinition

Aufgrund der strategischen Entscheidung, eine BI-Lösung bei *AVANCIS* einzuführen, wurden abteilungsübergreifende Arbeitskreise mit der Funktion, eine Ist-Analyse durchzuführen sowie Defizite und Schwachstellen aufzudecken, gebildet. Eine weitere Aufgabe bestand darin, Forderungen aufzustellen und Ziele zu definieren, deren Erreichung die festgestellten Mängel beseitigen sollten:

- Dynamische Standardberichte und Cockpits für betriebswirtschaftliche Kennzahlen
- Dynamische Standardberichte zur performanten Auswertung von Daten, welche innerhalb der Produktion von Solarmodulen generiert werden und zur Prozessverfolgung dienen
- Analysemodelle für Daten der operativen Systeme
- Möglichkeit zur Integration verschiedener operativer Datenquellen

Die aufgeführten Anforderungen lassen sich mit Hilfe von Business Intelligence abdecken, so dass im weiteren Verlauf näher das eingesetzte System *STAS CONTROL* beleuchtet wird.

3.2 Ist- und Sollanalyse

Hauptaktivitäten der Projektanalysephase waren die Ist- und Soll-Analyse sowie die Dokumentation der daraus resultierenden Anforderungen an ein Business Intelligence System in einem Lastenheft. Im Folgenden werden dazu die operativen Systeme, SAP ERP ECC 5.0 und Siemens *SILOC*, erläutert sowie die Schwerpunkte des Lastenheftes genannt und zusammengefasst. Dazu wird zunächst in einem kurzen Exkurs auf die, für die Produktionsdaten relevanten, Herstellungsprozesse eingegangen.

AVANCIS Solarmodule haben eine Größe von 158,3 x 66,0 cm und durchlaufen derzeit in der Fabrik 1 bis zur Fertigstellung 24 Prozessstationen (PID). Den Raum zwischen den Stationen überbrücken die Solarmodule in *MGVs* (Manuel Guided Vehicle) und werden aus diesen über einen *FETA* (Front End Tool Automation) automatisch der jeweiligen PID zur Bearbeitung zugeführt. Nach Abschluss des Teilprozesses wird dem *MGV* wiederum das Modul über einen *FETA* übergeben und kann an die nächste PID gefahren werden. In diesem Prozess fallen eine Vielzahl von Daten an den Prozessstationen, *MGVs* und *FETAs* an, welche in Datentelegramme geschrieben werden. Diese Telegramme dienen auf Initiative der Quelle zur Übertragung von Daten von der Quelle zu einem oder mehreren Empfängern.¹

Der Empfänger der Datentelegramme ist in diesem Zusammenhang die Siemens-Softwarelösung *SILOC*, welche eine wichtige Säule in der IT-Landschaft von *AVANCIS* darstellt. Das System hat die Aufgabe, Daten, welche im Herstellungsprozess der Module anfallen, in einer *ORACLE-Datenbank*² zu speichern und diese detaillierten Daten bedarfsorientiert zu präsentieren. Weiterhin werden in *Intermediate Documents* Daten an das *Enterprise-Resource-Planning-System* übermittelt, so dass *SILOC* ein Bindeglied zwischen Maschinensteuerungen und dem ERP-System darstellt und daher für die vertikale Integration der Systeme verantwortlich ist.

Als ERP-Lösung wurde bei *AVANCIS* im April 2008 die Standardsoftware der *SAP AG*, *SAP ERP ECC 5.0*, produktiv gesetzt. Diese Systemkategorie unterstützt mit einer

¹ Vgl. [Göh10]

² Für die Speicherung und Bereitstellung der Daten wird die *ORACLE-Datenbank* 10g verwendet.

konsistenten Datenbasis die betriebswirtschaftlichen Prozesse im Unternehmen.¹

Den Einführungsprozess begleitete mit der *applied international informatics GmbH (aii)* ein externer Dienstleister. Mit den Branchenlösungen *SPEED* bietet *aii* vorkonfigurierte Systeme, welche auf die Bedürfnisse der adressierten Branchen zugeschnitten sind und damit den Einführungsaufwand minimieren sollen. Eine für die Solarindustrie entwickelte Lösung stellt *aii* mit *SPEED Solar*, welche bei *AVANCIS* Anwendung findet.

Weiterhin übernimmt *aii* im Rahmen des Outsourcings Dienstleistungen wie etwa die Bereitstellung von Applikationslogik und Datenbankkapazität. Die dem ERP-System zugrundeliegende Datenbank arbeitet diesbezüglich mit dem relationalen Datenbankmanagementsystem *SAP MaxDB*.

SAP ERP bietet einen an der Organisationsstruktur der Unternehmen orientierten Aufbau. Die bei *AVANCIS* aktiv genutzten SAP-Module FI/CO², HCM³, MM⁴, PM⁵, PP⁶, QM⁷ und SD/CS⁸ lassen sich dem ECC⁹ 5.0 zuordnen und sind damit fundamentaler Bestandteil jedes SAP ERP 5.0.

Die Daten der operativen Quellen bilden die Basis für die zu erstellenden Auswertungen, welche im Lastenheft spezifiziert sind. Dabei wird im Lastenheft zwischen funktionalen Anforderungen an das BI-System und Anforderungen an den Auftragnehmer bezüglich Berichts- und Analysebereitstellung differenziert.

Zu den funktionalen Anforderungen an das einzuführende BI-System zählen insbesondere die Möglichkeiten, parametergesteuerte Reports zu erzeugen sowie innerhalb von multidimensionalen Datenmodellen Analysen durchzuführen. Dabei müssen vornehmlich Reports intuitiv erstellt werden können und darüber hinaus die Möglichkeiten eines Exports in verschiedene Formate wie etwa XLS, XLSX¹⁰ oder PDF¹¹ gegeben sein.

1 Vgl. [Fri08, S. 1]

2 Externes und internes Rechnungswesen

3 Personalwirtschaft

4 Materialwirtschaft

5 Instandhaltung

6 Produktionsplanung und -steuerung

7 Qualitätsmanagement

8 Vertrieb

9 ERP Central Component

10 Excel Spreadsheet

11 Portable Document Format

Weiterhin sollen Berichte dieser Formate zeitgesteuert im Hintergrund generiert und damit die Verteilung etwa per E-Mail durchführbar oder die Ablage in Netzwerkordnern möglich sein. Diese Methoden der Verteilung haben den Vorteil, dass sie zum Zeitpunkt der Einsichtnahme des Berichtes keinen Zugriff auf die Datenbasis erfordern.

Um die Durchführbarkeit im Hintergrund zu ermöglichen, muss die Selektion etwa eines Zeitraums dynamisch¹ erfolgen. Hingegen sollen aus einem BI-Portal heraus ausgeführte Berichte durch die Übergabe von ausgewählten Parametern wie Zeitraum, Gültigkeitsbereich oder etwa Buchungskreis an das Informationsbedürfnis des Benutzers angepasst werden. Die Auswahl der Parameter soll dabei in separaten Selektionsmasken erfolgen.

Das einzuführende BI-System muss in der Lage sein, verschiedene operative Datenquellen wie Datenbanken und Dateien mit den Formaten XLS, XLSX und CSV zu verarbeiten. Primäres Ziel dabei ist es, Datenbanken mit den Datenbankmanagementsystemen ORACLE 10g für die Abbildung von produktionsbezogenen Prozessdaten und die SAP-Datenbank MaxDB für die Abbildung von kaufmännischen Kennzahlen einzubinden. Strukturänderungen an operativen Datenquellen wirken sich gegebenenfalls auf den Informationsbedarf des Managements aus. Die daraus resultierenden notwendigen Strukturänderungen innerhalb der dispositiven Systeme sollen für die interne IT mit möglichst geringem Aufwand umsetzbar sein.

AVANCIS hält zum serverseitigen Betrieb der Softwarekomponenten einen Server mit dem Betriebssystem Windows Server 2008 und dem Datenbankmanagementsystem *Microsoft SQL Server* bereit. Daraus leitet sich die Forderung nach einer Kompatibilität der zu installierenden BI-Komponenten mit diesen Plattformen ab. Weiterhin ist es notwendig, dass die Clients der BI-Lösung auf PCs mit dem Betriebssystem Windows XP lauffähig sind. Bei der Verwendung von Thin-Clients ist die Unterstützung des MS Internet Explorers nötig, da dieser im Unternehmen als Standardbrowser verwendet wird.

Gegenstand des Lastenheftes sind weiterhin Anforderungen bezüglich Standardberichten und Analysen, die bei Inbetriebnahme des Systems vorliegen müssen. Im Lastenheft ist eine Priorisierung der produktionsbezogenen Themen festgehalten, so dass Reporting und

¹ Mit Hilfe eines dynamisch generierten Zeitraums können etwa Berichte für den vorangegangenen Tag, die aktuelle Woche oder das laufende Jahr erzeugt werden.

Analysen von produktionsbezogenen Prozessdaten Vorrang genießen. Es sollen unter anderem tagesbezogene Produktionsmengen, aber auch detaillierte Messwerte zur Auswertung verfügbar sein.

Produktionsberichte. Der Produktionsbericht soll für den definierten Zeitraum aggregierte Werte wie Produktionsmengen nach Leistungsklassen, Ausbuchungsmengen mit Gründen oder Durchlaufzeiten in Tabellen oder Diagrammen darstellen. Für die Verteilung per E-Mail oder über Netzwerkordner sollen die Berichte automatisch im Hintergrund generiert und anschließend in das PDF-Format exportiert werden. Die automatische Erzeugung des Produktionsberichts muss dabei jeweils für den vorangegangenen Tag für die verschiedenen Zeitfenster Tag, Woche und Monat erfolgen. Weiterhin müssen zur manuellen Auswahl des Zeitfensters parametergesteuerte Produktionsberichte zum Abruf aus einem Portal vorliegen.

Modulbezogene Kritische Parameter. Innerhalb des Produktionsprozesses von Solarmodulen fallen maschinen- und artikelbezogene Messwerte an. Durch *AVANCIS* werden diesbezüglich KPIs, also prozesskritische Parameter, definiert. Die Messwerte sollen jeweils in einem Diagramm entlang einer Zeitachse visualisiert werden. Weiterhin ist, wie in Grafik 15 zu sehen, die numerische messwertspezifische Darstellung des Mittelwertes, der Stichprobengröße, der Standardabweichung, des Maximal- und Minimalwertes sowie des Medians vorgesehen. Im Vorfeld der Abbildung ist es möglich, mit Hilfe einer Selektionsmaske einen Zeitraum und die darzustellenden Messwerte zu definieren.

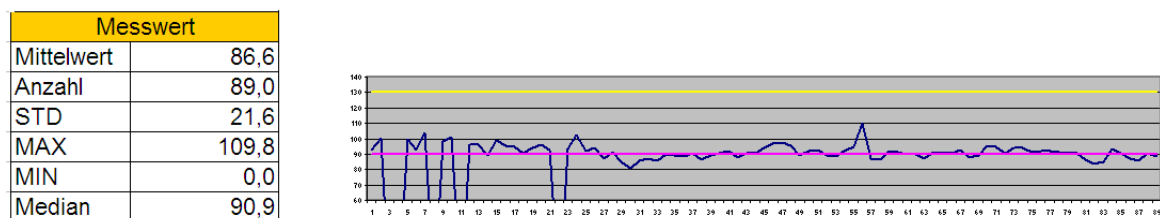


Abbildung 15: Entwurf zur Darstellung von prozesskritischen Parametern

DMC-Bericht. Jedem Modul wird innerhalb des Produktionsprozesses ein einmalig vergebener Data-Matrix-Code zugewiesen. Messwerte die während des

Fertigungsprozesses für ein DMC anfallen werden in eine ORACLE-Datenbank geschrieben. Auf Basis dieser Daten wird ein Bericht gefordert, der ausgewählte Messwerte modulbezogen in einer Übersicht tabellarisch darstellt. Die Auswahl des Moduls erfolgt in einer vorgelagerten Selektionsmaske, welche eine Suchfunktion für DMCs unter der Verwendung von Wildcards¹ anbietet.

Bericht Ausbuchen. Bei der Produktion von Solarmodulen fallen durch Prozessabweichungen Module an, welche die Marktreife nicht erreichen bzw. nicht erreichen können und daher den Fertigungsprozess vorzeitig verlassen. Diese Bewegungsart liegt in einer Tabelle des Produktionsleitsystems vor, so dass auf dieser Basis ein Bericht zur tabellarischen Darstellung von Modulen mit Ausbuchungsort und -gründen vorgesehen ist.

Bericht Verpackungseinheiten. Im Anschluss an den Fertigungsprozess werden die Solarmodule in Einheiten von jeweils 20 Modulen verpackt. Mit Hilfe einer Selektionsmaske kann eine Verpackungseinheit gesucht² und anschließend als Filter zur Ausführung des Reports gesetzt werden. Die numerisch-tabellarische Darstellung soll Angaben zu den Solarmodulen der Verpackungseinheit wie etwa Leistungskennzahlen aufweisen.

Darüber hinaus werden durch die *AVANCIS GmbH & Co. KG* Auswertungen und Analysen von kaufmännischen Daten gefordert. Als Datenquelle dient diesbezüglich das operative System SAP ERP 5.0, welches durch den Outsourcing-Partner *aii* betreut wird. Die Reports und Analysen halten als performante Auswertungsmöglichkeit betriebswirtschaftliche Kennzahlen wie etwa Umsatz, Lagerbestand oder Auftragseingänge bereit. Der potentielle Auftragnehmer hat weiterhin Erfahrung mit der Verarbeitung von SAP-Daten nachzuweisen und stellt zur Auswertung eines breiten Spektrums an betriebswirtschaftlichen Kennzahlen vordefinierte Berichte und Analysen zur Verfügung. Dieser Leistungsumfang ist in einem Angebot darzulegen.

Weiterhin soll der Auftragnehmer während des Implementierungsprozesses in die Lage versetzt werden, Datenquellen eigenständig zu erschließen, Berechnungen durchzuführen

1 Wildcards sind Platzhalter wie „%“ oder „*“ für andere Zeichen.

2 Durch die Verwendung von Fragmenten der Verpackungseinheitsnummer und Wildcards als Suchkriterium.

Vergleichbarkeit zu gewährleisten.

3.3 Design & Implementierung

3.3.1 Datenintegration

Zur Sicherstellung der Versorgung des *Data Warehouse* mit entscheidungsrelevanten Daten werden aufgrund von zunehmender Komplexität des ETL-Prozesses¹ Werkzeuge zur *Filterung*, *Harmonisierung*, *Aggregation* und *Anreicherung* herangezogen, so dass die Entwicklung von unternehmenseigenen Schnittstellen in den Hintergrund tritt.²

Mit *Talend Open Studio* ist es möglich, mit Hilfe einer Oberfläche Quellcode in *Pearl* oder *Java* durch die zielorientierte Zusammenstellung von Elementen zu generieren und diesen anschließend zeit- oder ereignisgesteuert ausführen zu lassen.

Talend ist mit dem Produkt *Talend Open Studio* Marktführer im Segment Open Source ETL-Werkzeuge.³ Um den Marktanteil weiter auszubauen und neue Finanzierungsquellen für das Projekt zu erschließen, wurde das OEM-Partnerprogramm *powered by Talend* initiiert. In diesem Zusammenhang bezeichnet Original Equipment Manufacturer (OEM) die Lizenzsoftware Dritter, z. B. *Talend Open Studio*, welche durch Softwareanbieter mit eigenen Produkten ergänzt und in einem Paket vermarktet wird.

Talend konnte bisher zahlreiche Unternehmen für das Partnerprogramm, wie etwa die *STAS GmbH* mit Sitz in Reilingen, gewinnen.⁴ Durch die Teilnahme an *powered by Talend* ist die *STAS GmbH* befugt, eine im Erscheinungsbild auf das Unternehmen abgestimmte Version von *Talend Open Studio* unter dem Namen *STAS CONTROL ETL powered by Talend* am Markt anzubieten. Weiterhin sind Erweiterungen zu *Talend Open Studio* innerhalb von *STAS CONTROL ETL* zur effizienten und damit schnelleren Durchführung von ETL-Projekten auf die *STAS GmbH* zurückzuführen.

1 Siehe Kapitel „ETL“

2 Vgl. [Sch00, S. 125]

3 Vgl. [Jas10]

4 Vgl. [Tal10]

Damit handelt es sich bei *STAS CONTROL ETL* um eine Softwarelösung, die zu 100 Prozent auf *Talend Open Studio* basiert, im Rahmen von *powered by Talend* durch die *STAS GmbH* vermarktet und zur effizienten Gestaltung von ETL-Prozessen durch Entwickler der *STAS GmbH* erweitert und optimiert wird.

Der Einstieg in *STAS CONTROL ETL* beginnt mit dem Anlegen eines Projektes und der Einrichtung des Repositorys. Beim Anlegen des Projektes muss weiterhin die Entscheidung fallen, ob der zu generierende Code in der Programmiersprache Pearl oder Java vorliegen soll. Um die Erweiterungen der Firma *STAS* zu nutzen, ist es jedoch erforderlich, das Projekt in *Java* umzusetzen.

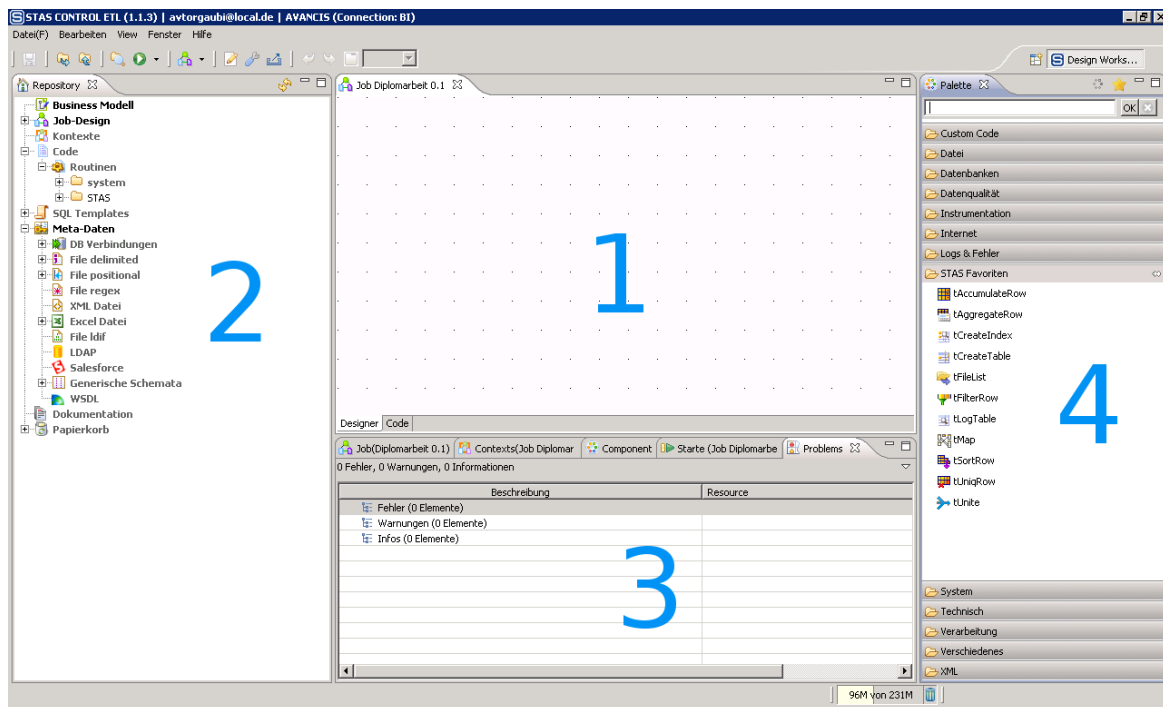


Abbildung 17: GUI STAS CONTROL ETL

Im Anschluss an die Initiierung eines neuen bzw. an das Öffnen eines bereits vorhandenen Projektes wird die Oberfläche von *STAS CONTROL ETL* sichtbar, die aus folgenden vier Bereichen besteht:

1. Arbeitsbereich

2. Repository
3. Konfiguration und Steuerung
4. Palette

Dem Repository kommt innerhalb von *STAS CONTROL ETL* eine herausragende Bedeutung zu, da in diesem sämtliche Elemente, die für ein Projekt generiert wurden, abgelegt sind und bei Bedarf aufgerufen und modifiziert werden können. Die wichtigsten Kategorien im *Repository* sind dabei das *Job Design* sowie die Kategorie der *Metadaten*.

Das Element *Job Design* bildet alle Datenflüsse des generierten Projektes ab und bietet damit eine grafische und funktionale Darstellung des gesamten ETL-Prozesses vom Datenimport über die verschiedenen Arten der Transformation bis zum Datenexport in das Data Warehouse. Die Abbildung der Prozesse erfolgt im Arbeitsbereich durch die zugehörigen Komponenten der Palette.

Während die Kategorie des *Job Designs* aktive Metadaten enthält, liegen in der Kategorie Metadaten in *STAS CONTROL ETL* semiaktive Metadaten vor.¹ In dem Cluster Metadaten können Strukturen von Ziel- und Quelltabellen definiert werden, wobei verschiedene Datenbanksysteme wie *MSSQL*, *MySQL* oder *ORACLE* sowie Dateiformate wie CSV, XLS, TXT und XML oder auch Anwendungsprotokolle wie LDAP² unterstützt werden. Entscheidende Erleichterung für die Implementierung eines ETL-Prozesses mit *STAS CONTROL ETL* birgt die Möglichkeit des Drag and Drop, so dass definierte Quell- oder Zielstrukturen in den Arbeitsbereich eines Jobs hineingezogen werden können. Bei der Erweiterung etwa einer Exceltabelle um eine Spalte, ist lediglich die definierte Struktur im Repository anzupassen, da im Anschluss ein Update der betroffenen Jobs durchgeführt wird.

Im Repository sind weiterhin die Kategorien Code, SQL-Templates, Business Modells und Kontexte hinterlegt. Hervorzuheben ist in diesem Zusammenhang das Cluster mit der Bezeichnung Code, durch welches auf einfachem Weg Erweiterungen in *STAS CONTROL ETL* integriert werden können. Diese Möglichkeit wird auch durch die Java-Entwickler der

¹ Siehe Kapitel Metadaten

² Mit der Integration von LDAP bzw. Lightweight Directory Access Protocol ist es möglich, Benutzerdaten zum Beispiel aus dem Active Directory auszulesen.

STAS GmbH genutzt, so dass *STAS CONTROL ETL* 130 Funktionen bereithält, welche nicht Bestandteil des *Talend Open Studio* sind und damit die Implementierung eines ETL-Prozesses in Bezug auf Berechnungen, Konvertierungen sowie den Umgang mit zeitbezogenen Daten und Zeichenketten erleichtern.

Der linke Bereich der Oberfläche von *STAS CONTROL ETL* zeigt nach dem Starten eines Projektes die Palette welche thematisch geordnete Elemente zur Gestaltung des ETL-Prozesses im Job Design enthält, die per *Drag and Drop* in den Arbeitsbereich gezogen und anschließend im Segment Konfiguration und Steuerung bearbeitet werden können.

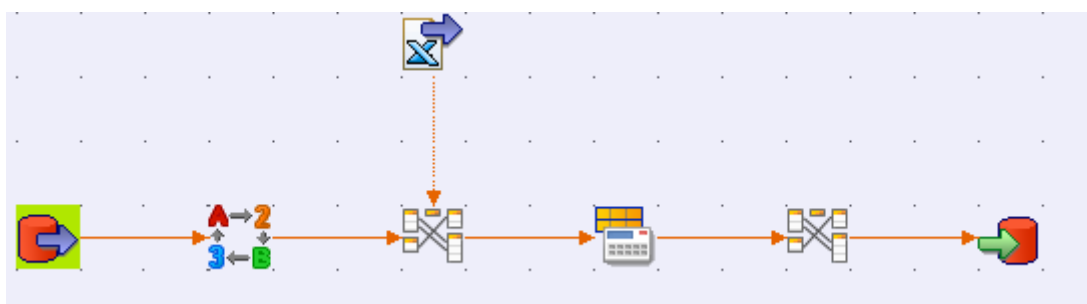


Abbildung 18: Prozessbeispiel in STAS CONTROL ETL

Die grafische Darstellung zur Umsetzung des durch die Elemente und Verbindungspfeile definierten Algorithmus erfolgt im Arbeitsbereich, so dass die Schritte zum Extrahieren, Transformieren und Laden sukzessive durchgeführt werden. Die Pfeile haben zusätzlich die Funktion, den Datenfluss an das nachstehende Element zu übergeben.

Im Anschluss an die Jobmodellierung kann der generierte Quellcode exportiert werden, so dass der Java- bzw. Pearl-Quellcodes unabhängig von *STAS CONTROL ETL* mittels einer Batchdatei angestoßen werden kann.

Mit *STAS CONTROL ETL* werden bei *AVANCIS* die Datenquellen der operativen Systeme SAP ERP ECC 5.0, Siemens *SILOC* sowie Dateien mit strukturierten Daten erschlossen. Zur Aufbereitung der Daten, welche in der SAP-Datenbank *MaxDB* vorliegen, konnte die *STAS GmbH* mit Hilfe des Unternehmens *COGIMO GmbH* einen vordefinierten ETL-Prozess anbieten, so dass lediglich eine partielle Prozessanpassung erfolgen musste. Die Generierung von Modellprozessen ist daher bei Standardsoftware mit hohem Marktanteil

wie etwa für die verschiedenen SAP ERP Releases¹ sinnvoll.

Im Zuge der Einführung einer BI-Lösung bei *AVANCIS* stand die *STAS GmbH* dabei erstmals vor der Herausforderung, die Datenbasis von Siemens *SILOC* aufzubereiten, so dass diesbezüglich nicht auf einen vorgefertigten ETL-Prozess zurückgegriffen werden konnte. Daraus ergab sich die Notwendigkeit, einen individuellen ETL-Prozess zu generieren. Die detaillierte Darstellung dieser Abläufe ist im Rahmen der Diplomarbeit nicht möglich, jedoch soll anhand des Beispiels aus Kapitel 3.2 exemplarisch ein Teilprozess des gesamten ETL-Prozesses beschrieben werden. Wie in Abbildung 19 aufgezeigt, liegen die für diese Auswertung benötigten Daten in der *ORACLE*-Datenbank des Siemens *SILOC* vor.

Filterung. Zur Berechnung der Ausbuchungsmengen wird über eine Datenbankschnittstelle eine Verbindung zum Siemens *SILOC* hergestellt und die Struktur der benötigten Tabellen im Repository des ETL-Werkzeugs angelegt und gespeichert. Diesbezüglich werden die Tabellenstrukturen der *EQ_VALUE* und der *MOVEMENT* zur Erfassung der Ausbuchungsmengen und die Struktur der Tabelle *DAY_REPORT* zur Erhebung der Produktionsmengen im Repository benötigt. Anschließend ist zur Erstellung einer logischen Abfolge ein Job in *STAS CONTROL ETL* zu generieren, welcher zu Erschließung der Datenquellen auf die definierten Metadaten zurückgreifen kann.

Harmonisierung. Mit der Harmonisierung erfolgt die Vereinigung² der *ORACLE*-Tabellen *EQ_VALUE* und *MOVEMENT*, um einen konsistenten Datenbestand in der Zieltabelle zu erhalten. Damit Konsistenz bzw. Widerspruchsfreiheit hergestellt werden kann, wird als Primärschlüssel die Modulnummer herangezogen. Die Tabelle *DAY_REPORT* enthält bereits auf Tagesbasis voraggregierte Produktionswerte, wobei der Aufbau dieser Tabelle sich von den Tabellen *EQ_VALUE* und *MOVEMENT* unterscheidet.

Ausbuchung			
Modulnummer	Equipment	Grund	Datum
AVANCIS0101...1	EQ_02	Kontaktieren	10.03.10
AVANCIS0101...2	EQ_03	Prozessfehler	11.03.10
AVANCIS0101...3	Verpackung	Kratzer	12.03.10

Produktionsmengen					
Datum	EQ_02	EQ_03	EQ_04A	Verpackung
10.03.10	500	495	490		480
11.03.10	501	496	491		481
12.03.10	502	497	492		482

Abbildung 19: Schematischer Aufbau der Tabellen Ausbuchungen (*EQ_VALUE* und *MOVEMENT*) und Produktionsmengen (*DAY_REPORT*)

- ¹ SAP verzeichnete im Jahr 2004 einen Marktanteil von 56 Prozent im Bereich Anwendungssoftware für Unternehmen. Vgl.: [Foc10]
- ² Im Sinne des SQL-UNION Befehls: [SQL-Anweisung 1] UNION [SQL-Anweisung 2]

Um im weiteren Verlauf eine Relation zwischen den Tabellen herstellen zu können, muss die Struktur der beiden Tabellen aufeinander abgestimmt werden. Dazu wird die Tabelle *DAY_REPORT* transformiert, indem die einzelnen Spalten unter Verwendung des Datums in eine Tabellenzeile geschrieben werden.

Datum	EQ_02	EQ_03	EQ_04A	Verpackung
10.03.10	500	495	490		480
11.03.10	501	496	491		481
12.03.10	502	497	492		482

Datum	Equipment	Anzahl
10.03.10	EQ_02	500
10.03.10	EQ_03	495
10.03.10	Verpackung	480

Abbildung 20: Transformation der Tabelle *DAY_REPORT*

Aggregation. Im Vorfeld dieser Stufe muss die Relevanz von detaillierten Daten geklärt sein, so dass die Aggregation auf eine bestimmte Verdichtungsstufe erfolgen kann. Aufgrund der Vorverdichtung der Produktionsmengen auf Tagesbasis, richtet sich der Focus innerhalb dieses Teilschrittes auf den Datenstrom der Ausbuchungen. In diesem Zusammenhang kann die Verdichtung auf eine Ebene erfolgen, in welcher die Attribute Datum, Equipment (EQ) und Ausbuchungsgrund¹ die detaillierteste Stufe darstellen. Der Datenstrom wird bezüglich dieser Werte gruppiert und die Modulnummern auf Basis dieser Gruppierungen aufgezählt.

Datum	Equipment	Grund	Anzahl
10.03.10	EQ_02	Kontaktieren	1
...

Abbildung 21: Aggregation der Ausbuchungsmengen auf die Ebene Datum, Equipment, Grund

¹ Im Sinne des SQL-Befehls GROUP BY

Anreicherung. Dem Teilprozess der Anreicherung kann bereits das Hinzufügen der Spalte *Anzahl* und das Aufzählen der Module unter Beachtung der genannten Attribute zugewiesen werden. Um den Inhalt der Datenströme *Ausbuchungen* und *Produktionsmengen* im Data Warehouse wiederzufinden, werden in dieser Datenbank die Zieltabellen eingerichtet. Das Anlegen der Zieltabellen wird dabei ebenfalls mit *STAS CONTROL ETL* realisiert, wobei die Strukturen der Tabellen im Anschluss in das Repository von *STAS CONTROL ETL* aufgenommen werden. Damit ist es möglich, per Drag and Drop die Zieltabellen in den Job zu integrieren und die in den Prozessen *Filterung*, *Harmonisierung* und *Aggregation* aufbereiteten Daten durch das Anstoßen des Jobs in die Tabellen zu schreiben.

Die Aktualität der Ausbuchungsdaten im DWH soll viertelstündlich überprüft und so die Ausbuchungsdaten möglichst aktuell im Data Warehouse vorliegen, wobei die Kapazität des operativen Systems Siemens *SILOC* beachtet werden muss. Um diese Herausforderung zu bewältigen, wurde folgender Weg gewählt:

1. Auslesen des Datensatzes im DWH mit dem maximalen Zeiteintrag
2. Auslesen der Datensätze aus der Tabelle *MOVEMENT* und *EQ_VALUE*, die einen Zeiteintrag besitzen, welcher größer als der maximale Zeiteintrag der DWH-Tabelle ist

Die Produktionsmengen werden hingegen jeweils nach Mitternacht für den Vortag aktualisiert und können aufgrund der geringen Datenmenge¹ nach dem Löschen der Zieltabelle in das DWH geschrieben werden.

Auf Basis der im DWH vorliegenden Daten kann anschließend die Erstellung des OLAP-Würfels und des Reports in *IBM Cognos 8.4* erfolgen.

3.3.2 Analyse und Reporting

Ist die Befüllung des DWH mit aktuellen sowie entscheidungsrelevanten Daten

¹ Jeweils ein Datensatz pro Tag

sichergestellt, bietet *IBM* mit *Cognos 8.4* eine Standardsoftware, mit welcher *Ad-Hoc-Abfragen*, *Ad-Hoc-Analysen*, *Standardberichte*, *Dashboards* oder *Cockpits* generiert werden können. Zur Erstellung und Präsentation dieser Auswertungen werden über das Webportal *IBM Cognos Connection* auf die Anwendung zugeschnittene Werkzeuge angeboten. Darüber hinaus bietet *Cognos 8.4* Möglichkeiten zur Modellierung von relationalen und multidimensionalen Datenmodellen an.

Um den Lieferumfang und die Funktionen von *Cognos 8.4* besser einschätzen zu können, lohnt ein Blick auf die Architektur der BI-Software. Wie in der Abbildung 3 deutlich wird, setzt *Cognos* dabei auf eine mehrstufige Architektur der Ebenen Datenhaltung, Anwendung, Webserver und Benutzerschnittstelle.

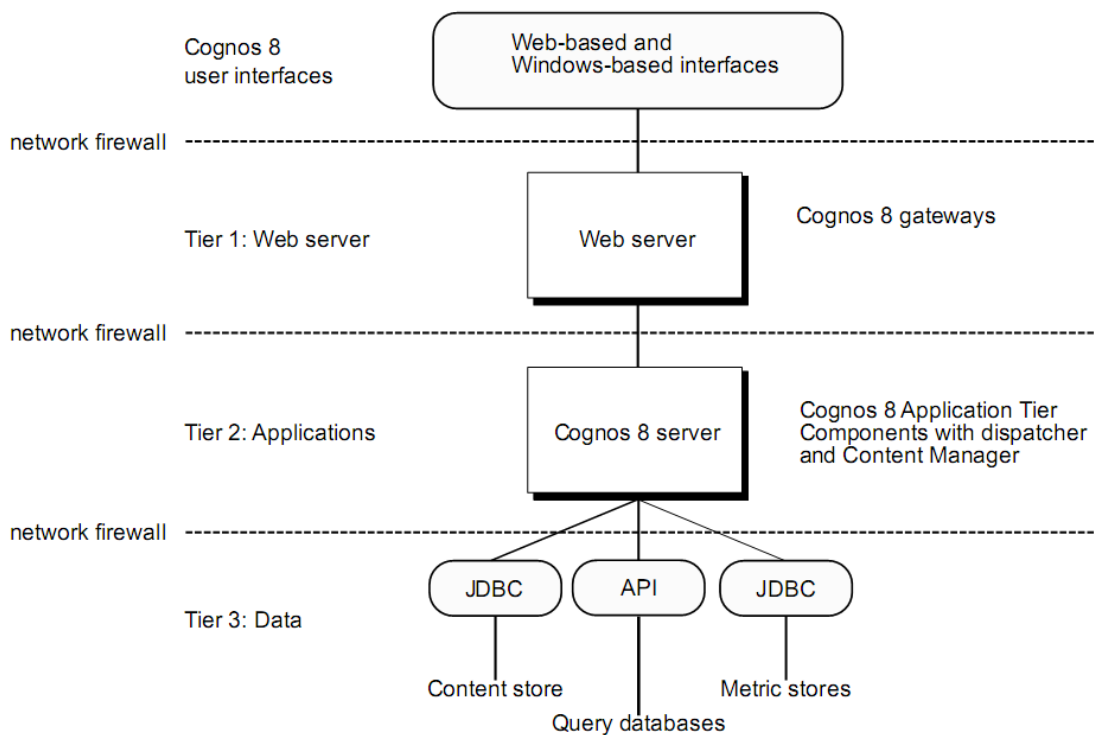


Abbildung 22: logische Cognosarchitektur. [IBM10]

Innerhalb der Schicht *Cognos 8 User Interface* wird zwischen web- und windowsbasierenden Benutzerschnittstellen unterschieden. Die windowsbasierten Anwendungen wie der *Frameworkmanager*, der *Metric Designer* und der *Transformer* dienen der Modellierung von Datenmodellen und sind, wie auch die webbasierten

Anwendungen, Gegenstand im weiteren Verlauf des aktuellen Kapitels. Kennzeichnend für Web-Anwendungen ist die Verwendung eines Browsers wie z. B. den *Microsoft Internet Explorer* zur Präsentation von Inhalten. Die Darstellung erfolgt in, auf die Datenbasis und die Bedürfnisse der Benutzer abgestimmten BI-Tools, die sich nach der Anmeldung im Webportal *Cognos Connection* ausführen lassen.

Die Schicht Webserver untergliedert sich in die Elemente Firewall und *Cognos 8 Gateway*, wobei die Firewall das erste Security-Level zwischen dem Frontend und dem *Cognos 8 Gateway* darstellt, so dass die komplette Kommunikation bezüglich Abfragen und deren Ergebnissen zwischen User Interface und Gateways überwacht wird. *Cognos 8.4* unterstützt dabei verschiedene Arten von Gateways bzw. Web-Server-Schnittstellen wie *CGI*, *ISAPI*, *Apache_mod* und *Servlet*.

Abfragen der Benutzerschnittstellen gelangen über den Webserver zur Applikationsschicht und werden innerhalb dieser Stufe bearbeitet. Bestandteile des *Cognos 8*-Servers sind dabei der *Content Manager*, verantwortlich für *Cognos 8*-Inhalte und der *Dispatcher*, welcher Dienste verwaltet und Abfragen an freie Kapazitäten weiterleitet.

Integraler Bestandteil des *Content Manager* zur Authentifizierung von Benutzern ist der *Access Manager*, welcher Zugangsregeln auf Basis des *Authentication Store* erstellt. Weiterhin speichert der *Content Manager* semiaktive Metadaten in einer relationalen Datenbank, die definierte Standardberichte, Packages¹, Server Konfigurationen oder persönliche Benutzereinstellungen enthält und unter dem Namen *Content Store* fungiert.

Eine weitere Komponente des Webserver stellt der *Dispatcher* dar, der sämtliche konfigurierte sowie aktivierte Dienste verwaltet. Dazu gehören unter anderem Präsentationsdienste für das Webportal *Cognos Connection* oder die zeitgesteuerte Erstellung von Berichten.

Die Schicht der Datenhaltung bildet die unterste Ebene der Cognosarchitektur. Diese relationale Datenbank dient der Bereitstellung von Packagedefinitionen, Reportergebnissen, Konfigurationen und Verzeichnisstrukturen und ist kein Bestandteil von *Cognos 8.4*, weshalb Cognos auf die Bereitstellung eines *MSSQL*-Servers, einer *IBM DB2* oder einer *ORACLE*-Datenbank angewiesen ist.

1 Metadatenmodell

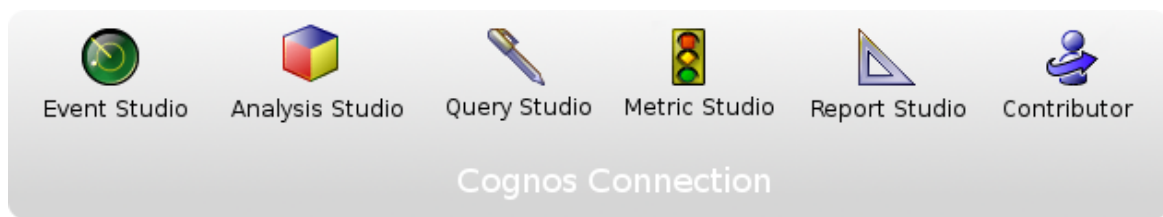


Abbildung 23: Cognos Business Intelligence Tools

Damit dem Management entscheidungsrelevante Informationen bereitgestellt werden können, erfolgt nach der Erzeugung einer dispositiven Datenbasis und der Installation von *Cognos 8.4* die Modellierung von Metadaten unter der windowsbasierten Oberfläche Frameworkmanager. Die Metadatenmodelle werden innerhalb von *Cognos 8.4* unter dem Begriff *Packages* zusammengefasst und haben die Aufgabe, Stamm- und Bewegungsdaten aus einer oder mehreren Quellen logisch sowie strukturell zu organisieren¹, so dass Informationen zugeschnitten auf die Bedürfnisse von Personen oder Personengruppen bereitgehalten werden. Weiterhin können Metadaten bezüglich Zugriffsbeschränkungen und Sprachzugehörigkeit Bestandteil eines Packages sein und so auch multilinguale Business Intelligence Projekte unterstützen. Im Anschluss an die Initiierung bzw. an die Modifizierung wird das Package im *Content Store* publiziert und im *IBM Cognos Connection* zur Verfügung gestellt, um auf relationalen Daten Berichte im *Report Studio* zu erstellen oder Ad-hoc-Analysen im *Query Studio* durchzuführen.

Um multidimensionale Datenmodelle für die Auswertungswerkzeuge des Cognos Webportals bereitzustellen, werden Inhalte des Packages hingegen in eine *Impromptu Query Definition* (IQD) geschrieben. Dateien mit der Namenserweiterung *IQD* lassen sich beispielsweise durch das Cognos-Werkzeug *Transformer* öffnen, so dass die im Package definierten Datenquellen als Grundlage zur Modellierung eines *PowerCubes*² herangezogen werden können. Daher dient *IBM Cognos Transformer* als administratives Werkzeug, das den Anwender in die Lage versetzt, auf Basis der in *IQDs* definierten relationalen Daten multidimensionale Datenmodelle zu erstellen. Die grafische Benutzeroberfläche des *Transformer* bietet dazu Möglichkeiten, Attribute der relationalen Datenbasis als Dimensionen zu deklarieren und so betriebswirtschaftliche Kennzahlen aus

1 Aufbau von Relationen zwischen Stamm- und Bewegungsdaten

2 Bezeichnung innerhalb von IBM Cognos für multidimensionale Datenmodelle

verschiedenen Perspektiven zu betrachten.

Um den *PowerCube* Endanwendern bereitzustellen, ist es erforderlich, diesen zu generieren und das Datenmodell unabhängig von *IBM Cognos Transformer* verfügbar zu machen. Dazu werden Dimensionen und Kennzahlen in einer *MDC*-Datei angelegt, welche anschließend über den *Frameworkmanager* als *Package* publiziert wird und damit im *Content Store* als Datenbasis zur Verfügung steht. Damit die Aktualität der Kennzahlen und Dimensionen sichergestellt wird, kann die Erzeugung des Würfels zeitgesteuert über eine *Batchdatei*¹ angestoßen werden.

Das Webportal *Cognos Connection* versteht sich als Ausgangspunkt zur Ausführung der im Lieferumfang von *Cognos 8.4* enthaltenen webbasierten Studios, mit denen der Zugriff auf relationale wie auch dimensionale Daten des *Content Stores* ermöglicht wird. „Mit dem *Query Studio*, dem *Report Studio*, dem *Event Studio*, dem *Metric Studio* und dem *Analysis Studio* lassen sich Ad-hoc-Berichte, detaillierte Berichte und datenbasierte Benachrichtigungen erstellen sowie Daten überwachen und analysieren.“²

Die über das Webportal zugänglichen Studios beziehen Daten aus relationalen oder mehrdimensionalen Datenquellen. Zur inhaltlichen und strukturellen Abgrenzung werden Datenquellen eines im *Frameworkmanager* erstellten *Packages* den Auswertungstools zugewiesen. Auf Basis der Metadaten des gewählten *Packages* erfolgt die Erstellung von Berichten, Analysen und Abfragen innerhalb des Webportals.

Mit dem *Query Studio* stellt *Cognos 8.4* ein Tool zur Generierung von Ad-Hoc-Abfragen auf Listenbasis zur Verfügung, mit welchem Daten aus relationalen wie auch multidimensionalen Quellen angezeigt, gefiltert, sortiert und formatiert werden können. Das *Query Studio* orientiert sich dabei am Bedienkonzept von *Microsoft Office*-Produkten, so dass Benutzer mit wenig Schulungsaufwand in die Lage versetzt werden sollen, einfache Listenabfragen zu erstellen. Die generierten Abfragen können im Anschluss abgespeichert werden und stehen damit über die Verzeichnisstruktur auch anderen Nutzern zur Verfügung.

Das *Analysis Studio* ist ein webbasierter Analyse-Client, mit dessen Hilfe

¹ Datei zur Durchführung von sequentiellen und nicht-interaktiven Aufgaben.

² [Vol09, S. 68]

multidimensionale Daten über verschiedene Analysetechniken wie *Drill-Down*, *Roll-Up*, *Split* oder etwa *Merge* untersucht werden können und damit Hypothesen verifizierbar sind. Innerhalb der Benutzerschnittstelle stellen sich Kennzahlen und Dimensionen in einem Baumdiagramm dar und können intuitiv per *Drag and Drop* in den Arbeitsbereich, welcher aus einer Kreuztabelle besteht, hineingezogen werden. Weiterhin stehen Möglichkeiten zur Formatierung, zum Sortieren oder etwa zur Visualisierung von Werten durch Diagramme zur Verfügung.

Das *Report Studio* beinhaltet Funktionen, welche die Erstellung sowohl einfacher Berichte mit Listenabfragen, als auch komplexer Applikationen wie zum Beispiel Dashboards oder Cockpits auf Basis von multidimensionalen wie auch relationalen Daten unterstützen. Der Funktionsumfang dieser webbasierten Anwendung schließt dabei unter anderem Diagramme, Listen, Kreuztabellen, Landkarten und verschiedene Varianten zur optischen Gestaltung ein. Mit der Verwendung von Drop-Down-Auswahllisten oder Zeitauswahlfenstern können Berichte darüber hinaus dynamisch gestaltet werden.

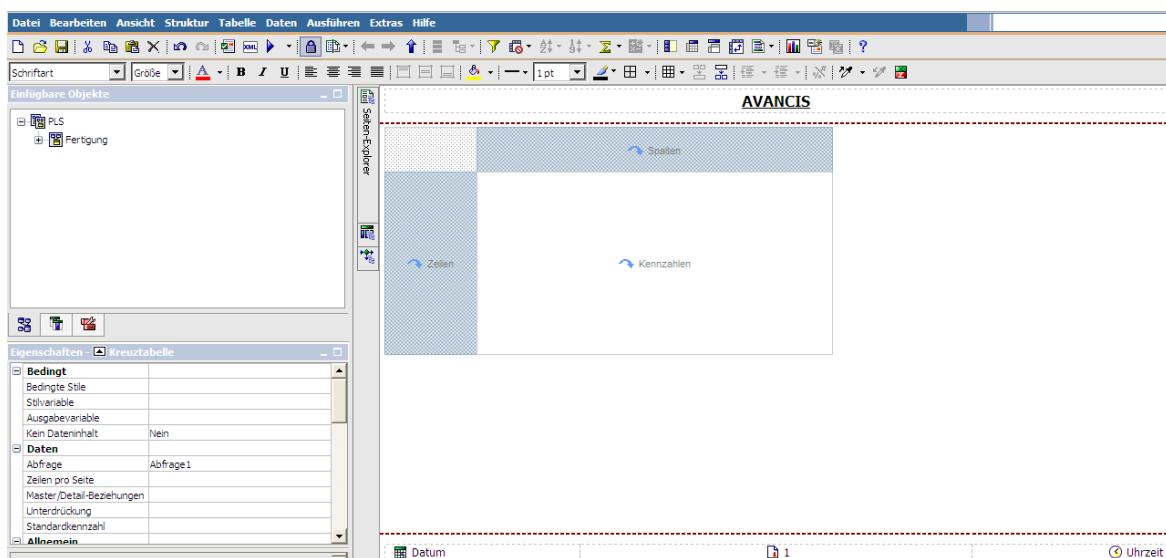


Abbildung 24: GUI des Cognos 8 Report Studio

Die Erstellung der Berichte erfolgt auf Basis eines Packages in der GUI des *Report Studio*, das sich aus der Arbeitsfläche und den Segmenten *Einfügbare Objekte* sowie *Eigenschaften* zusammensetzt. Innerhalb des *Report Studios* wird bezüglich der

Einfügbaren Objekte zwischen *Quellen*, *Datenelementen* und *Berichtselementen* unterschieden.

Quellen. Die Quellen enthalten alle relationalen oder multidimensionalen Daten des ausgewählten Packages.

Datenelemente. Unter diesem Reiter befinden sich alle Datenelemente der Abfragen, die im aktuellen Report verwendet werden. Ein Bericht kann dabei n Abfragen und diese Abfragen wiederum n Datenelemente enthalten. Auf Basis der Gestaltung der Abfragen wird jeweils nach Art des Packages¹ ein *SQL*²- oder *MDX*³-Statement erzeugt und bei Ausführung auf die Datenquelle abgesetzt.

Berichtselemente. Diese Elemente sind Objekte wie Tabellen oder Textelemente zur Berichtsgestaltung sowie Listen oder Kreuztabellen zur strukturierten Wiedergabe von Datenelementen.

Alle einfügbaren Objekte können dabei per Drag and Drop in den Arbeitsbereich des Seiten-Explorers gezogen werden, so dass eine schnelle und effiziente Berichterstellung möglich ist. Darüber hinaus kann innerhalb der Sichten Abfragen-Explorer und Bedingungs-Explorer mit Berechnungen und Filtern die Ausgabe der Abfragen manipuliert und durch die Deklaration von Bedingungen etwa betriebswirtschaftliche Kennzahlen bei Verletzung einer Grenze in der Ausgabe farblich hervorgehoben werden. Die Ausgabe der mit Report Studio erstellten Berichte kann dabei in verschiedenen Formaten wie HTML, PDF, XLS, XLSX oder CSV erfolgen, wobei das Standardformat HTML zur Darstellung im Browser dient und die Ausgabe eines Berichts im PDF-Format sich hervorragend zur zeitgesteuerten Verteilung per E-Mail eignet.

In den bisher behandelten Studios geht die Initiative zur Untersuchung von betriebswirtschaftlichen Kennzahlen vom Anwender aus, indem er mit dem *Report Studio* erstellte Standardberichte, Cockpits oder Dashboards abrufen, Analysen mit dem *Analysis Studio* durchführt oder Ad-Hoc-Abfragen mit dem *Query Studio* generiert. Durch eine andere Arbeitsweise ist das *Event Studio* gekennzeichnet. Mit diesem Werkzeug ist es

1 Multidimensional oder relational

2 Structured Query Language

3 Multidimensional Expressions

möglich, Schwellwerte für eine Kennzahl zu hinterlegen, welche durch die Initiierung eines Agenten überwacht wird. Stellt der Agent eine Verletzung der Grenzwerte fest, können Aktionen wie E-Mail-Versand oder die Implementierung einer Zusatzinformation ins Portal durchgeführt werden. Mit dem *Metric Studio* steht in diesem Zusammenhang ein weiteres Werkzeug zur Verfügung, mit dem Kriterien deklariert und anschließend in einer Scorecardumgebung überwacht werden können. Innerhalb einer *Scorecard* werden verschiedene Überwachungselemente bzw. Metriken definiert und durch strukturelle und grafische Elemente visualisiert. Metriken bilden dabei den Ist-Wert, den Soll-Wert und die Differenz beider Werte ab.

Um die vorgestellten Funktionalitäten von *Cognos 8.4* zu verdeutlichen, wird an dieser Stelle erneut das Beispiel aus Kapitel 3.2 aufgegriffen. Durch die Implementierung eines ETL-Prozesses zur Bereitstellung der Ausbuchungsdaten sowie der maschinenbezogenen Produktionsmengen wird durch das DWH eine aktuelle, konsistente Datenbasis zur Definition eines OLAP-Würfels sowie zur Erstellung eines Reports bereitgehalten.

Um die Tabellen des DWH in *Cognos 8.4* nutzen zu können, werden die Strukturen im *Frameworkmanager* eingelesen und als Metadaten gespeichert. Weiterhin ist es innerhalb dieses Werkzeugs möglich und in diesem Beispiel notwendig, Beziehungen zwischen verschiedenen Tabellen zu deklarieren. Um innerhalb der Nutzerschnittstellen Auswertungen zur Verfügung zu stellen, welche den prozentualen Anteil der Ausbuchungen an der Produktionsmenge angeben, müssen Beziehungen zwischen den Tabellen, wie in Abbildung 25 aufgezeigt, hergestellt werden.

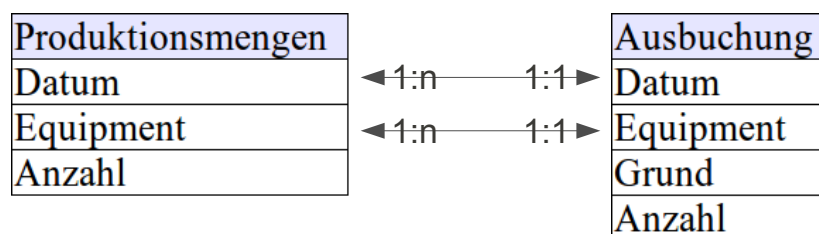


Abbildung 25: Deklarierte Beziehungen zwischen den Spalten der DWH-Tabellen

Die anschließende Bereitstellung der relationalen Daten mit Hilfe eines *Packages* zur

Generierung eines Reports und der Export der Metadaten in eine *Impromptu Query Definition* als Voraussetzung für die Erstellung eines OLAP-Würfels wurde eingangs beschrieben.

Zur Definition der Dimensionen für das multidimensionale Datenmodell wird innerhalb von *Cognos 8.4* das Werkzeug Transformer verwendet, dessen Arbeitsbereiche in der Abbildung 26 dargestellt werden. Nach dem Einlesen der *IQD* als Datenquellen können diese Spalten als Dimensionen oder als Kennzahlen deklariert werden.

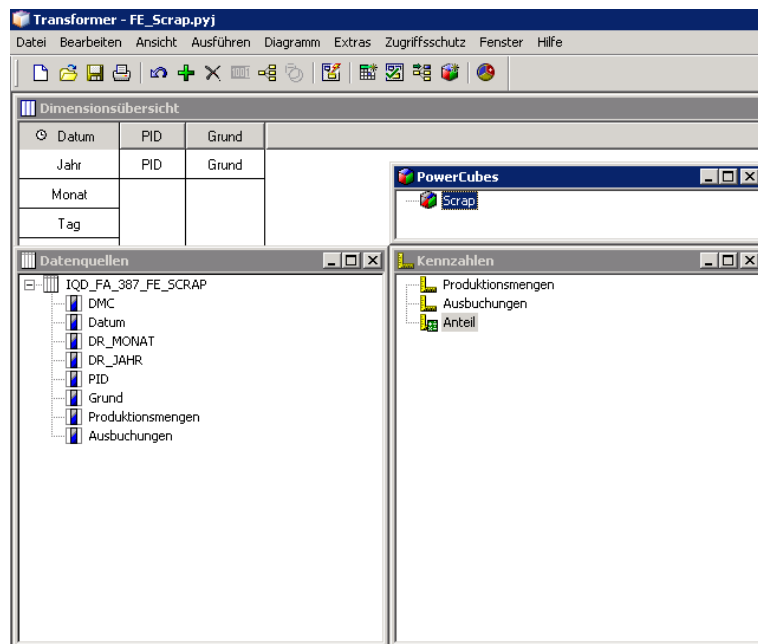


Abbildung 26: Modellierung eines OLAP-Würfels zur Darstellung der Ausbuchungswerte

Im Beispiel werden als Dimension das *Datum* mit den Hierarchien Jahr, Monat, Tag sowie die Dimensionen *Equipment* und *Ausbuchungsgrund* definiert. Als Kennzahlen dienen die *Produktionsmengen*, die *Anzahl der Ausbuchungen* sowie die berechnete Kennzahl *Anteil*, welche sich aus Ausbuchungen und Produktionsmengen ergibt.

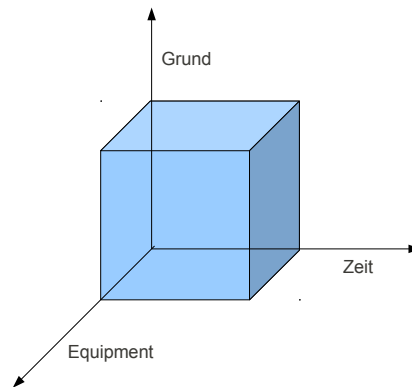


Abbildung 27: Dimensionen des generierten OLAP-Würfels

Im Anschluss an die Definition der Kennzahlen und Dimensionen kann das multidimensionale Datenmodell generiert und, wie in diesem Kapitel eingangs abgehandelt, über den *Frameworkmanager* im *Content Store* bereitgestellt werden. Damit können Analysen auf Basis dieses Würfels im webbasierten Werkzeug *Analysis Studio* durchgeführt und die verschiedenen Techniken zur Untersuchung eines multidimensionalen Datenbestands angewandt werden.

Weiterhin liegt mit einem Package ein relationales Datenmodell im *Content Store* vor, auf dessen Basis ein parametergesteuerter Bericht mit Hilfe des *Report Studios* erstellt werden kann, welches in Bezug auf das Beispiel verschiedene Möglichkeiten anbietet, die Daten visuell aufzubereiten. Die folgenden Abbildungen zeigen dazu die innerhalb eines Berichtes generierten Darstellungen in Form des *Trendcharts* und des *Paretodiagramms*. Um diese Diagramme dynamisch zu gestalten und im benötigten Rahmen Analysen zu ermöglichen, können mit Hilfe der Auswahlfelder *Zeitraum*, *Prozessstationen* und *Gründe* eingegrenzt werden.

Das Trendchart stellt dabei tagesbezogen die Anzahl der aufgetretenen Fehler an den jeweiligen Prozessstationen sowie die Relation zwischen Ausbuchungsmenge und Tagesproduktion dar.

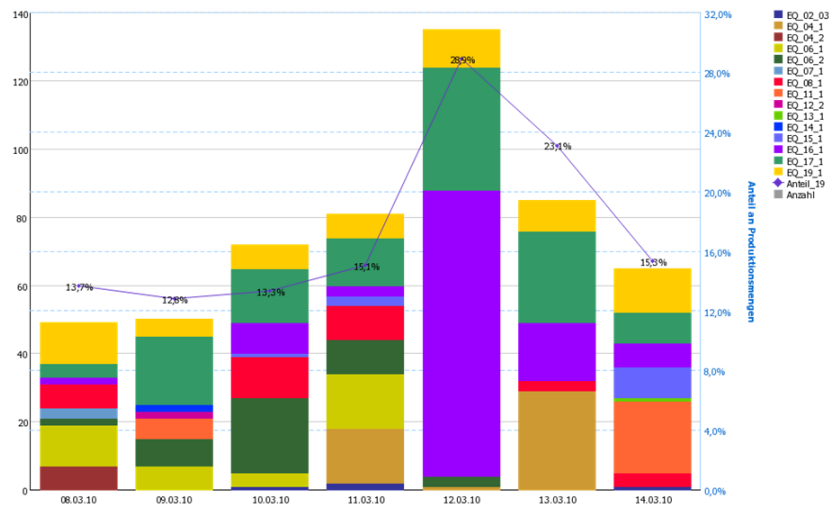


Abbildung 28: Trendchart

Das Paretdiagramm hingegen zeigt die Anzahl der Ausbuchungen kumuliert über einen ausgewählten Zeitraum unter Einbezug von Prozessstationen und Ausbuchungsgründen an.

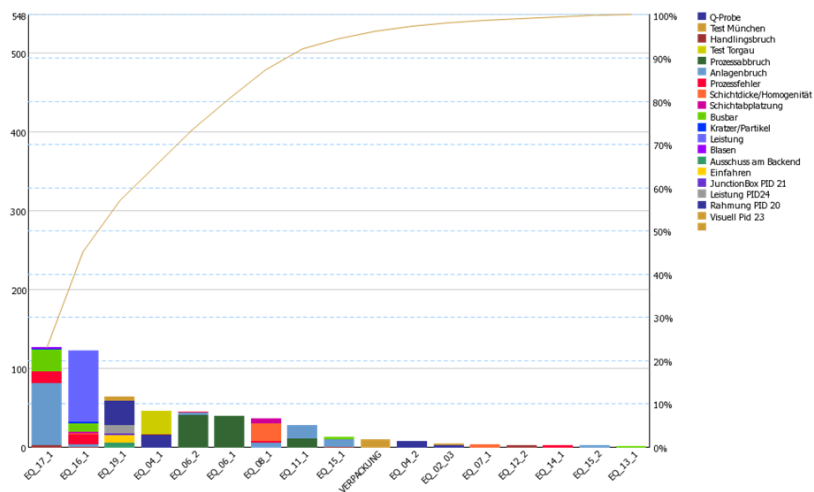


Abbildung 29: Paretdiagramm

IBM Cognos 8.4 bietet eine breite Palette an Werkzeugen zur Aufbereitung und Darstellung von Unternehmensdaten mit dem Ziel, Führungskräfte mit fundierten Informationen zu versorgen und damit die Entscheidungsqualität zu steigern. Weiterhin zeichnet sich Cognos 8.4 durch die intuitive Bedienbarkeit und die Vielfältigkeit des Webportals aus und ist somit für die Einführung einer BI-Lösung in einem

mittelständischen Unternehmen hervorragend geeignet. Diese Vielfalt kann jedoch im Rahmen einer Diplomarbeit lediglich in Fragmenten dargestellt werden.

3.3.3 Bereitstellung

Gegenstand der vorangegangenen Kapitel war die Generierung von managementunterstützenden Informationen mit *STAS CONTROL ETL* powered by talend und IBM Cognos 8.4. Wichtiger Bestandteil eines jeden BI-Projektes ist neben der Erstellung dieser Informationen ein Konzept zur Distribution von Inhalten des BI-Systems.

Dabei ist zunächst abzuklären, welchen Benutzern bzw. Benutzergruppen Zugang zum BI-System gewährt werden soll. Bei *AVANCIS* werden Daten zu Benutzern und deren Gruppenzugehörigkeit innerhalb des Verzeichnisdienstes Active Directory erstellt und gepflegt. Diese Datenbestände¹ können mit Hilfe der webbasierte Administrationsoberfläche von Cognos Connection eingelesen und zur Erstellung von cognosspezifischen Zugangsrechten verwendet werden. Diese Zugangsregeln werden im Anschluss an die Implementierung durch den Authentication Store bereitgestellt und limitieren damit den benutzerspezifischen Zugriff auf Inhalte des Content Manager.

In Kapitel 2.5 wurde diesbezüglich bereits auf die Differenzierung zwischen Zugriffsrechten auf spezifische Daten und Zugriffsrechten für spezifische Anwendungen eingegangen. Aufgrund der verschiedenen Bedürfnisse, Aufgabenbereiche und der an Business-Intelligence-Applikationen gerichteten Erwartungen wird die Verfügbarkeit der Cognos Anwendungen bzw. Studios für den einzelnen Nutzer durch standardisierte Benutzerlizenzen geregelt, welche nach dem *Named-User-Lizenzmodell*² vergeben werden. Eine Übersicht der unterschiedlichen Lizenzen und die damit verbundenen Zugriffsrechte werden in der Tabelle 3 aufgezeigt. Aufgrund der fehlenden Zugriffsfreigabe für das Cognos Webportal stellt der Lizenztyp *Remote Recipient* eine Besonderheit dar. Dieser Lizenztyp ist für Benutzer vorgesehen, welche sich nicht am System anmelden, dennoch

1 Die möglichen Quellen im Überblick: Active Directory Services , IBM Cognos Series 7, LDAP, NTLM , SAP.

2 Die Software kann ausschließlich durch registrierte, namentlich eingetragene (named) User genutzt werden.

via E-Mail von ereignis- oder zeitgesteuerten Auswertungen profitieren.

Lizenz		Remote Recipient	Recipient	Consumer	Business Author	Professional Author	Business Analyst	Professional	Administrator
Web	Cognos Connection	x	x	x	x	x	x	x	x
	Query Studio				x	x		x	x
	Report Studio					x		x	x
	Event Studio							x	x
	Metric Studio							x	x
	Analysis Studio						x	x	x
Windows	Metric Designer							x	x
	Frameworkmanger								x
	Transformer								x

Tabelle 3: Lizenztypen Cognos 8.4

Neben der Begrenzung der Zugriffsrechte auf spezifische Anwendungen muss weiterhin sichergestellt sein, dass ausschließlich die für den jeweiligen Aufgaben- und Verantwortungsbereich erforderlichen Daten bereitgestellt werden und verfügbar sind. Dazu werden in *Cognos 8.4* verschiedenen Möglichkeiten angeboten:

1. mit dem Frameworkmanager kann der benutzer- bzw. benutzergruppenspezifische Zugang über Webanwendungen auf bestimmte relationale Daten beschränkt werden
2. mit dem Werkzeug Transformer kann der benutzer- bzw. benutzergruppenspezifische Zugang über Webanwendungen auf multidimensionale Daten beschränkt werden

3. mit Hilfe der Administrationsschnittstelle kann eine Zugangsbeschränkung etwa auf einzelne Reports implementiert werden
4. Reports können mit anwenderbezogenen Filtern¹ erstellt werden

Der Zugang zu den windowsbasierten, administrativen Anwendungen ist wie in Tabelle 3 aufgezeigt, Nutzern mit zugewiesener Administrator- bzw. Professional-Lizenz vorbehalten. Als Mittel zur Verteilung von dispositiven Inhalten dienen hingegen die webbasierten Werkzeuge, zu denen die Benutzer über das Webportal Cognos Connection Zugang erhalten.

Mit den aufgezeigten Möglichkeiten können im praktischen Einsatz wie etwa bei *AVANCIS* die Zugriffsmöglichkeiten auf die Unternehmensdaten mit Hilfe von Cognos vielfältig eingeschränkt und vergeben werden. Durch die Administration des BI-Systems ist dabei zu prüfen, ob Zugangsrechte global, für Benutzergruppen oder für einzelne Benutzer vergeben werden können.

Das bei *AVANCIS* implementierte Berechtigungssystem ist an Abbildung 14 angelehnt, wobei Führungskräfte der oberen Ebene Zugriff auf verdichtete Werte aller Abteilungen erhalten, die zur strategischen Ausrichtung des Unternehmens dienlich sind. So kann etwa der am Markt erzielte Preis pro Watt ausschlaggebend für weitere bzw. zukünftige Entscheidungen des Topmanagements und damit von Interesse und Bedeutung sein.

Um den Aufwand bei Pflege und Erweiterung der Berechtigungsstrukturen so gering wie möglich zu halten, ist die in Abbildung 14 aufgezeigte Unterscheidung zwischen mittlerer und unterer Führungsebene nicht vorgenommen sondern eine Zusammenfassung dieser durchgeführt worden. Im Segment der mittleren und unteren Führungsebene sind dabei neben abteilungsspezifischen auch abteilungsübergreifende Daten zur Verfügung zu stellen, welche für integrierte Auswertungen essentiell sind. Bei *AVANCIS* muss etwa für Mitarbeiter des Vertriebs die Möglichkeit bestehen, auf Daten der Materialwirtschaft zuzugreifen, um etwa Informationen zum aktuellen verkaufsfähigen Lagerbestand zu erhalten.

¹ In die Datenabfrage wird ein Filter im Report Studio zur Überprüfung einer oder mehrerer Benutzername integriert. Diese Methode ist zu vermeiden, da diese im Widerspruch zu einer zentralen Berechtigungsimplementierung steht.

Zur Verteilung der dispositiven Informationen auf die verschiedenen Führungsebenen und Mitarbeiter bietet *IBM Cognos 8.4* die Möglichkeit über das BI-Portal auf die webbasierten Anwendungen zuzugreifen, wobei eine permanente Verbindung mit dem *Cognos 8*-Webserver bestehen muss. Um dem Anspruch gerecht zu werden Informationen jederzeit und von jedem Ort abrufen zu können, bietet *Cognos 8.4* die Möglichkeit Reports zeit- bzw. ereignisgesteuert auszuführen und die Ergebnisse per E-Mail in verschiedenen Ausgabeformaten zu versenden.

Bei *AVANCIS* wird auf diese Weise, wie im Lastenheft gefordert, eine Zusammenfassung der tagesbezogenen Produktionswerte an die zuständigen Mitarbeiter verteilt, so dass diese auch bei Abwesenheit etwa aufgrund von Dienstreisen mit aktuellen Informationen versorgt werden. Mittelfristig könnten auch kritische Statusmeldungen von einzelnen Prozessstationen an die verantwortlichen Ingenieure versandt werden, damit diese bei instabilen Prozessen unverzüglich stabilisierende Maßnahmen ergreifen und so die fortlaufende Fertigung von *AVANCIS*-Solarmodulen sicherstellen können.

Neben der zeit- und ereignisgesteuerten Verteilung von Informationen bietet *Cognos 8.4* mit *Cognos 8 Go! Office*¹ ein zusätzliches Modul zur Integration von BI-Inhalten in die MS Office-Umgebung. *IBM Go! Office* ist dabei als Add-In lokal zu installieren und bietet im Anschluss die Möglichkeit Elemente von Reports wie Datentabellen, Diagramme etc. in die Office-Anwendungen wie *Word*, *Excel* oder *Powerpoint* zu importieren und so zur anschließenden Bearbeitung zur Verfügung zu stellen. Mit dieser Methode können aktuelle Inhalte des BI-Systems in den genannten Microsoftprodukten dargestellt werden und Prozessschritte etwa zur Erstellung von Präsentationsfolien vereinfacht werden.

Bei *AVANCIS* kann auf diese Weise die Erstellung von *Powerpoint*-Folien für das Reporting an den Mutterkonzern teilweise automatisiert und damit das Top-Management und die im Erstellungsprozess einbezogenen Mitarbeiter entlastet werden.

Mit dem Webportal *IBM Cognos Connection* bietet *IBM Cognos 8.4* einen zentralen Zugang zu allen Informationen des BI-Systems, welche durch die aufgeführten Möglichkeiten ausschließlich dem adressierten Nutzerkreis zugänglich gemacht wird. Mit *Cognos 8 Go! Office* und der zeit- bzw. ereignisgesteuerten Verteilung kann die

1 Siehe <http://www-142.ibm.com/software/products/de/de/cognos-8-go-office/>

Bereitstellung von dispositiven Informationen auch ohne die Verbindung zur Datenbasis realisiert werden.

3.4 Produktivsetzung, Betrieb und Wartung

Die Produktivsetzung erfolgt idealerweise nach der Umsetzung des Gesamtkonzeptes und einer Testphase durch die Überführung eines Testsystems in ein Produktivsystem. Aufgrund des Zeitplans und der Priorisierung der produktionsnahen Themen war die Abweichung von den idealtypischen Einführungszyklen notwendig, damit einzelne Teilbereiche unverzüglich die Arbeit mit dem System aufnehmen konnten, um somit von den Ergebnissen des BI-Systems zu profitieren.

Die aufgestellten Anforderungen an das BI-System, welche im Abschnitt 3.2 thematisiert und im Lastenheft festgehalten sind, konnten durch die *STAS GmbH* erfüllt werden. Nach Abschluss der Einführung und der Bereitstellung erster Auswertungen ist nun der Betrieb und die Weiterentwicklung des Systems zu fokussieren, um eine hohe und breite Akzeptanz der Anwendung zu schaffen. Dies ist wesentlich für den Erfolg oder auch Misserfolg eines BI-Projektes.

Eine hohe Akzeptanz kann dann erreicht werden, wenn Anforderungen insbesondere zur Weiterentwicklung des Systems zeitnah umgesetzt werden können und den Anwendern des Unternehmens für diese Belange ein Ansprechpartner zur Verfügung steht. Dabei steht das Unternehmen vor der Entscheidung, ob Kompetenzen für Betrieb und Weiterentwicklung des Systems intern aufgebaut werden oder ein externer Partner für bestimmte Leistungen beauftragt wird. Die Möglichkeit, das BI-System innerhalb des Unternehmens zu betreuen und weiterzuentwickeln, hat dabei den Vorteil, dass Anforderungen zeitnah aufgenommen und umgesetzt werden können. Weiterhin sind Themen wie der Datenschutz und die Flexibilität bei Änderung der Anforderungen besser zu handhaben.

Für eine externe Lösung und damit für ein Outsourcing sprechen in vielen Fällen die geringeren zu erwartenden Kosten und eine höhere Transparenz dieser. Darüber hinaus zeichnen sich externe Dienstleister durch eine hohe Spezialisierung und damit durch einen

aktuellen Kenntnisstand aus.

Jedes Unternehmen muss ausgehend von seinen individuellen Ressourcen, Bedürfnissen und Ausrichtungen sowie den Kernkompetenzen und -aufgaben der internen IT eine für ihre Ziele passende Lösung finden, welche vor allem die Akzeptanz der Anwender sicherstellen muss.

4 Fazit & Perspektive

Mit der Lösung *STAS CONTROL* ist bei *AVANCIS* ein Business Intelligence System in Betrieb, welches mit IBM Cognos 8.4, *STAS CONTROL ETL* und den Metadaten für ein Data Warehouse alle erforderlichen BI-Komponenten bereithält. Die bei *AVANCIS* realisierte Struktur des BI-Systems ist im Folgenden grafisch dargestellt und zeigt schematisch die Kernaspekte der implementierten Lösung.

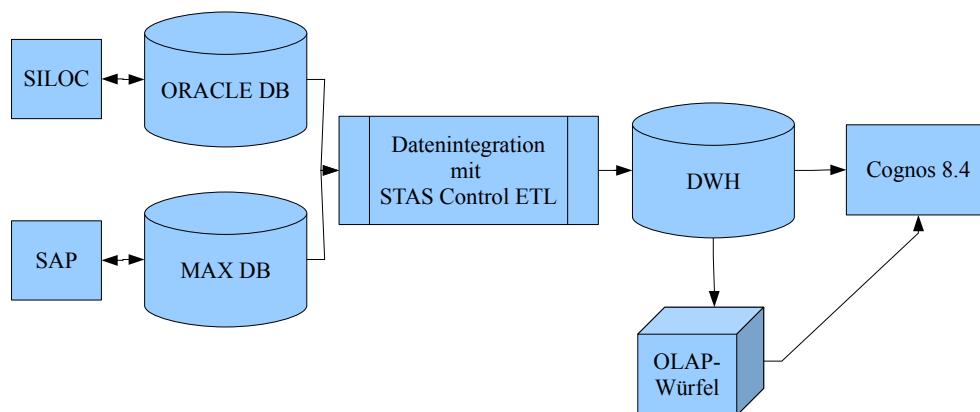


Abbildung 30: Implementierte BI-Struktur bei AVANCIS

Zum Aufbau dieser Struktur wurde mit der *STAS GmbH* ein Partner gewonnen, der vordefinierte Standardauswertungen beim Kunden implementiert, welche auf den bereits im Lieferumfang enthaltenen Data Warehouse Tabellen und OLAP-Würfeln basieren. Auf diese Weise wird eine Einführung einer schlüsselfertigen BI-Lösung innerhalb eines schmalen Zeitfensters ermöglicht, so dass die adressierten Mitarbeiter bereits nach einer kurzen Einführungsphase die Arbeit mit dem System aufnehmen können. Vordefinitionen werden durch die *STAS GmbH* für die Auswertung und Analyse von Standardkennzahlen wie Umsatz, Auftragseingänge oder etwa Materialbestand angeboten. Damit die Einführung innerhalb kürzester Zeit gelingt, ist es weiterhin erforderlich, dass eine standardisierte Berechnung der Kennzahlen erfolgen kann. Eine vom Standard abweichende Berechnung der Kennzahlen ist etwa dann notwendig, wenn durch tiefgreifendes Customizing des operativen Systems dieses vom Auslieferungszustand abweicht. In diesem Fall können die Ergebnisse des BI-Systems different zur Realität sein,

so dass diese analysiert und daraufhin die jeweiligen ETL-Prozesse angepasst werden müssen.

Vordefinierte Lösungen unterscheiden sich also in diversen Bereichen von individuell auf das jeweilige Unternehmen zugeschnittenen BI-Systemen. Die nachfolgende Tabelle erklärt diese Unterschiede im Einzelnen und benennt Vor- und Nachteile.

Attribut	Vorgefertigte BI-Lösung ¹	Individuallösung ²
Zeitraumen	Es ist von einer kurzen Einführungsphase auszugehen.	Hierbei handelt es sich um eine angepasste Lösung, daher ist i. d. R. eine längere Einführungsphase nötig.
Kosten	Es kann, bedingt durch die kürzere Einführungsphase, mit geringeren Kosten kalkuliert werden.	Im Vergleich mit der vorgefertigten Lösung sind höhere Kosten zu erwarten.
Erfolgsaussichten & Stabilität	Vorgefertigte Lösungen sind i. d. R. bereits bei anderen Unternehmen im Einsatz und dadurch umfangreich getestet. Auf dieses Know-How kann zurückgegriffen werden, so dass bei der Implementierung einer solchen BI-Lösung von hohen Erfolgsaussichten ausgegangen werden kann.	Die Erfolgsaussichten sowie die Stabilität der Anwendung sind abhängig vom vorhandenen Zeitrahmen und dem durch das Unternehmen bereitgestellten betriebswirtschaftlichen und informationstechnischen Fachwissen.
Anpassung auf die Unternehmensbedürfnisse	Der Fokus liegt bei vorgefertigten Lösungen insbesondere auf der Darstellung von Standardkennzahlen wie Umsatz etc. Die Einbeziehung von individuellen Kennzahlen wie etwa Materialbestand in Watt ist nur durch die Anpassung und somit durch Individualisierung der BI-	Bei der Individuallösung werden insbesondere Auswertungen bereitgestellt, welche im Lastenheft definiert wurden. Die Erstellung dieser erfolgt mit der Ausrichtung auf individualisierte Kennzahlen. Weiterhin werden in den operativen Systemen nicht gepflegte Attribute bzw. Kennzahlen vom Datenimport

¹ Wie etwa STAS Control

² Eine Individuallösung in diesem Sinne könnte etwa Talend Open Studio zur Datenintegration mit Cognos 8.4 zur Visualisierung der Inhalte darstellen. Vorgefertigte ETL-Prozesse und Auswertungen liegen dabei nicht vor.

	Lösung möglich.	ausgeschlossen.
Datenquellen	Vorgefertigte Lösungen sind meist auf <i>eine</i> Datenquelle zugeschnitten, wie z.B. die in einem ERP-System zugrunde liegende Datenbank.	Die individuelle Zusammenführung von mehreren Datenquellen im Rahmen des Initialprojektes ist möglich.
Corporate Identity (CI)	Optik und Design der Reports und des Webportals sind nicht auf die CI des Unternehmens zugeschnitten.	Webportal und Report werden individuell erstellt und können damit auf Basis der CI angepasst werden.
Projekt	Die Abteilungen werden oft erst nach der Präsentation erster Ergebnisse zur Ergebnisverifizierung, Fehlersuche und anschließender Ursachenforschung intensiv einbezogen.	Die Mitarbeiter werden stärker eingebunden. Einzelschritte und Teilergebnisse werden den adressierten Abteilungen präsentiert. Kritik und Verbesserungsvorschläge wirken sich so direkt auf das Endergebnis aus.
Weiterentwicklung	Das Know-How für die Lösung liegt vorwiegend beim Einführungspartner. Im Vorfeld von Anpassungen durch das Unternehmen selbst ist die intensive Auseinandersetzung mit den Prozessen und Strukturen des BI-Systems notwendig. Anpassungen und Erweiterung sind weiterhin durch externe Unternehmen möglich.	Das Unternehmen begleitet den Einführungsprozess intensiv oder führt diesen selbstständig durch. Damit sind Mitarbeiter des Unternehmens mit den Prozessen und Strukturen des BI-Systems betraut und können somit die Anpassung bzw. Erweiterung des Systems durchführen.

Tabelle 4: Gegenüberstellung von vorgefertigter und individueller BI-Lösung

Mit Hilfe der Software *STAS CONTROL* war es möglich, eine Business Intelligence Lösung im Rahmen vorher festgelegter zeitlicher Vorgaben, zum einen an die Bedürfnisse des Unternehmens angepasst und zum anderen unter Nutzung bereits bestehender, etablierter Lösungen, bei *AVANCIS* zu implementieren. Dabei wurde im Bereich der kaufmännischen Daten auf die vorgefertigte Lösung zurückgegriffen, da diese Daten im *ERP-System mySAP ERP 2005* vorlagen und die *STAS GmbH* Erfahrungen bei der Integration von SAP-Daten in ein BI-System nachweisen konnte. Die Integration von produktionsnahen Daten aus dem Siemens *SILOC* stellen im BI-System hingegen die individualisierte Variante dar und erforderten daher die unternehmensspezifische Erstellung von Strukturen zum Datenimport und die Erstellung von eigenen Reports.

Die Entscheidung für eine der genannten Varianten richtet sich nach den Bedürfnissen, Voraussetzungen, Möglichkeiten und Perspektiven des Unternehmens, welches die Einführung eine BI-Lösung anstrebt. Ausschlaggebend wirkt sich ebenfalls das Marktsegment aus, in dem sich das Unternehmen bewegt.

Wie eingangs erläutert, befinden sich der Markt für Photovoltaikprodukte in einer Phase entscheidender Veränderungen. Nur *die* Unternehmen, die sich an die sich verändernden globalen und nationalen Bedingungen anpassen können, werden gestärkt aus diesem schwierigen, jedoch auch herausfordernden Umfeld hervorgehen. Um zeitnah auf den flexiblen Markt reagieren zu können, ist es erforderlich, dass Führungskräften entscheidungsrelevante interne wie auch externe Informationen bereitgestellt werden und diese somit Entschlüsse fassen können, welche auf fundierten Informationen basieren.

Weiterhin ist das System für ein junges und aufstrebendes Unternehmen wie *AVANCIS* hervorragend geeignet. Durch die Erschließung neuer Standorte bzw. Errichtung neuer Produktionsstätten erscheint es möglich, dass neben der Modifikation der im ERP-System abgebildeten Unternehmensstruktur weitere operative Systeme den Unternehmensablauf dokumentieren und steuern. Eine BI-Lösung wie *STAS CONTROL* bietet dabei die nötige Flexibilität, so dass weitere Datenquellen integriert und basierend auf diesen Daten zusätzliche Auswertungen und Analysen erstellt werden können.

Mit der Einführung von *STAS CONTROL* hat *AVANCIS* eine zukunftsichernde Entscheidung getroffen. Die Entscheidungsträger können nun innerhalb kurzer Zeit mit

unternehmens- und produktionsrelevanten Informationen versorgt werden, um das Unternehmen trotz des sich schnell verändernden Marktes weiter zu etablieren und voranzubringen.

4.1 Integration von Data Mining

Mit *STAS CONTROL* befindet sich bei *AVANCIS* ein BI-System in der Anwendung, das über eine Vielzahl von Komponenten des engeren, weiteren und analyseorientierten BI-Verständnisses verfügt, jedoch ist derzeit keine Lösung im Einsatz, welche Data Mining ermöglicht. Um neben der Hypothesenverifizierung auch Hypothesen zu generieren, erscheint es dienlich, das bestehende BI-System um eine Data Mining-Komponente zu erweitern.

Im Moment wird die Data Mining-Lösung *Minitab* von einzelnen Abteilungen genutzt. Diese stellt eine autarke Softwarelösung dar, welche lokal installiert wird und deren Ergebnisse nur unter Nutzung von dritter Software¹ weiteren Mitarbeitern zugänglich gemacht werden kann. Eine weitere Stand-Alone-Lösung stellt die in Abschnitt 2.4 erläuterte Open Source-Software *RapidMiner* dar, welche im Gegensatz zu *Minitab* in dieser Variante kostenlos zur Verfügung gestellt wird.

Eine Stand-Alone-Lösung kann bei der vorliegenden IT-Infrastruktur jedoch nicht zielführend sein, so dass die Integration eines Data Mining Werkzeuges in eine Dreischichten-Architektur erforderlich ist. Das Tool *RapidMiner* bietet dabei die Möglichkeit, Data Mining innerhalb der Enterprise Developer Edition als Webanwendung einzusetzen. Der Quellcode dieser Anwendung liegt jedoch nur zum Teil offen und es sind Lizenzgebühren zu entrichten. Ein weiterer Nachteil besteht darin, dass diese Anwendung nur schwer in das bestehende BI-System *STAS CONTROL* und die dazugehörige Visualisierungskomponente Cognos 8.4 zu integrieren ist.

Ein ganzheitlicher Ansatz wäre hingegen die bereits vorhandene *Microsoft*-Software *Visual Studio* zu verwenden. Diese Applikation beinhaltet eine Komponente zum

¹ dritte Software im Sinne von MS Office-Anwendungen

automatisierten Data Mining, mit welcher Hypothesengenerierung betrieben werden kann. Die Bereitstellung der aufgestellten Hypothesen kann anschließend über *Cognos 8.4* innerhalb eines Reports erfolgen und somit ein Ausgangspunkt für die daran anschließende Hypothesenverifizierung sein.

III Literaturverzeichnis

- [Abt10] Abts, Dietmar; Mülder, Wilhelm (Hrsg.): Masterkurs Wirtschaftsinformatik. Kompakt, praxisnah, verständlich-12 Lern- und Arbeitsmodule, 1. Auflage, Wiesbaden, 2010.
- [Ale10] Alexander, Sascha; Vaske, Heinrich: Die Topanbieter im BI-Markt, 2009, URL: <http://www.computerwoche.de/top100/2009/software/1902863/>, (Abgerufen am: 22.03.2010).
- [Alp08] Alpar, Paul; Grob, Heinz Lothar; Weimann, Peter; Winter, Robert: Anwendungsorientierte Wirtschaftsinformatik - Strategische Planung, Entwicklung und Nutzung von Informations- und Kommunikationssystemen, 5. Auflage, Wiesbaden, 2008.
- [Alp00] Alpar, Paul; Niedereichholz, Joachim (Hrsg.): Data Mining im praktischen Einsatz. Verfahren und Anwendungsfälle für Marketing, Vertrieb, Controlling und Kundenunterstützung, 1. Auflage, Wiesbaden, 2000.
- [And10] Andrews, Chris: IBM Completes Acquisition of Cognos, 2008, URL: <http://www-03.ibm.com/press/us/en/pressrelease/23423.wss>, (Abgerufen am: 14.02.2010).
- [Bau04] Bauer, Andreas; Günzel, Holger (Hrsg.): Data Warehouse Systeme. Architektur, Entwicklung, Anwendung, 2. Auflage, Heidelberg, 2004.
- [Bei06] Beierle, Christoph; Kern-Isberner, Gabriele: Methoden wissensbasierter Systeme, 3. Auflage, Wiesbaden, 2006.
- [Cha06] Chamoni, Peter; Gluchowski, Peter (Hrsg.): Analytische Informationssysteme. Business Intelligencetechnologien und -Anwendungen, 3. Auflage, Berlin, Heidelberg, 2006.
- [Cha05] Chamoni, Peter; Gluchowski, Peter; Hahne, Michael: Business Information Warehouse. Perspektiven betrieblicher Informationsversorgung und Entscheidungsunterstützung auf der Basis von SAP-Systemen, 1. Auflage, Berlin, Heidelberg, 2005.
- [Clo10] Cloer, Thomas: Die zehn (vielleicht) wichtigsten IT-Persönlichkeiten der letzten 40 Jahre, URL: <http://www.tecchannel.de/job-karriere->

- seminar/483936/die_zehn_vielleicht_wichtigsten_it_persoenlichkeiten_der_letzten_40_jahre/index7.html, (Abgerufen am: 12.01.2010).
- [Dis03] Disterer, Georg; Fels, Friedrich; Hausotter, Andreas: Taschenbuch der Wirtschaftsinformatik, 2. Auflage, Leipzig, 2003.
- [Foc10] TOMORROW FOCUS Portal GmbH: SAP gewinnt weiter Marktanteile, URL: http://www.focus.de/finanzen/news/baerenstark_aid_87726.html, (Abgerufen am: 12.03.2010).
- [Fri08] Frick, Detlev; Gadatsch, Andreas; Schäffer-Külz, Ute: Grundkurs SAP ERP, 1. Auflage, Wiesbaden, 2008.
- [Gab09] Gabriel, Roland; Gluchowski, Peter; Pastwa, Alexander: Data Warehouse & Data Mining, 1. Auflage, Herdecke, Witten, 2009.
- [Gab03] Gabriel, Roland; Röhrs, Heinz-Peter: Gestaltung und Einsatz von Datenbanksystemen. Data Base Engineering und Datenbankarchitekturen., 1. Auflage, Berlin, Heidelberg, 2003.
- [Gad08] Gadatsch, Andreas: Grundkurs Geschäftsprozess-Management, 5. Auflage, Wiesbaden, 2008.
- [Gen10] Gentsch, Peter: Business Intelligence for better decisions, 2009, URL: <http://www.intelligence-group.com/downloads/BI-for-better-decisions.pdf>, (Abgerufen am: 12.01.2010).
- [Glu15] Gluchowski, Peter: Business Intelligence, In: HMD-Praxis der Wirtschaftsinformatik. 38. Jg., 2001, Nr. 222, S. 5-15.
- [Glu08] Gluchowski, Peter; Gabriel, Roland; Dittmar, Carsten: Management Support Systeme und Business Intelligence, 2. Auflage, Berlin, Heidelberg, 2008.
- [Göh10] Göhner, Peter: Automatisierungstechnik, URL: http://www.ias.uni-stuttgart.de/pa1/lehrmaterialien/uebungen/sl_can.pdf, (Abgerufen am: 12.11.2009).
- [Gra04] Grasl, Oliver; Rohr, Jürgen; Grasl, Tobias : Prozessorientiertes Projektmanagement, 1. Auflage, München, 2004.
- [Gro00] Grothe, Martin, Gentsch, Peter: Business Intelligence. Aus Informationen

- Wettbewerbsvorteile gewinnen, 1. Auflage, Bonn, 2000.
- [Han96] Hannig, Uwe (Hrsg.): Data Warehouse und Managementinformationssysteme, . Auflage, Berlin, Heidelberg, 1996.
- [Han02] Hannig, Uwe (Hrsg.): Knowledge Management und Business Intelligence, 1. Auflage, Berlin, Heidelberg, 2002.
- [Han09] Hansen, Hans Robert; Neumann, Gustaf: Wirtschaftsinformatik 1 – Grundlagen und Anwendungen, 10. Auflage, Stuttgart, 2009.
- [Hol99] Holthuis, Jan : Der Aufbau von Data Warehouse-Systemen: Konzeption - Datenmodellierung – Vorgehen, 2. Auflage, Wiesbaden, 1999.
- [Hut03] Huthmann, Andreas: Metadaten und Datenqualität in Data Warehouses, 1. Auflage, Norderstedt, 2003.
- [IBM10] International Business Machines Corp.: IBM® Cognos® 8 Business Intelligence 8.2 product documentation, URL: http://www-01.ibm.com/support/docview.wss?rs=3442&context=SS9RTN&dc=DA400&uid=swg27014542&loc=en_US&cs=UTF-8&lang=en&rss=ct3442db2#cognos8_bi8.2, (Abgerufen am: 05.11.2009).
- [Jas10] Jaspersoft Corporation: Jaspersoft und Talend arbeiten an Datenintegrationspack zusammen, dass für BI Kunden spezifisch angepasst wird, 2009, URL: <http://www.jaspersoft.com/de/node/988> , (Abgerufen am: 11.03.2010).
- [Jun00] Jung, Reinhard; Winter, Robert : Data Warehousing Strategie: Erfahrungen, Methoden, Visionen, 1. Auflage, Berlin Heidelberg, 2000.
- [Kem07] Kemper, Hans-Georg; Baars, Henning: Business Intelligence Arbeits- und Übungsbuch: Glossar, Aufgaben, Lösungsskizzen, 1. Auflage, Wiesbaden, 2007.
- [Kem04] Kemper, Hans-Georg; Mehanna, Walid; Unger, Carsten: Business Intelligence. Grundlagen und praktische Anwendungen, 1. Auflage, Wiesbaden, 2004.
- [Ker02] Kerner, Simone: Analytisches Customer Relationship Management in Kreditinstituten. Data Warehouse und Data Mining als Instrumente zur Kundenbindung im Privatkundengeschäft, 1. Auflage, Wiesbaden, 2002.
- [Kle06] Kletti, Jürgen (Hrsg.): MES-Manufacturing Execution System. Moderne

- Informationstechnologie zur Prozessfähigkeit der Wertschöpfung., 1. Auflage, Berlin, Heidelberg, 2006.
- [Knö06] Knöll, Heinz-Dieter; Schulz-Sacharow, Christoph; Zimpel, Michael: Unternehmensführung mit SAP ® BI, 1. Auflage, Wiesbaden, 2006.
- [Kol09] Kollmann, Tobias: E-Entrepreneurship, 3. Auflage, Wiesbaden, 2009.
- [Kön03] König, Wolfgang (Hrsg.): Taschenbuch der Wirtschaftsinformatik und Wirtschaftsmathematik, 2. Auflage, Frankfurt am Main, 2003.
- [Kre08] Krems, Burkhardt: Management-Ebenen: operativ - taktisch – strategisch, 2008, http://www.olev.de/o/operativ_usw.htm, 2008.
- [Kre09] Krems, Burkhardt: Management-Ebenen: operativ - taktisch – strategisch, 2008, URL: http://www.olev.de/o/operativ_usw.htm, (Abgerufen am: 12.11.2009).
- [met10] metaApplicationCenter GmbH: Metadaten im Data Warehouse, 2009, URL: http://www.metaapplicationcenter.de/fr_metaw.htm, (Abgerufen am: 15.02.2010).
- [Möh02] Möhlenbruch, Dirk; Hartmann, Michaela (Hrsg.): Der Handel im Informationszeitalter. Konzepte-Instrumente-Umsetzung., 1. Auflage, Wiesbaden, 2002.
- [Nav08] Navrade, Frank: Strategische Planung mit Data-Warehouse-Systemen, 1. Auflage, Wiesbaden, 2008.
- [Oeh06] Oehler, Karsten: Corporate Performance Management mit Business Intelligence Werkzeugen, 1. Auflage, München, Wien, 2006.
- [Oeh99] Oehler, Karsten: OLAP, 1. Auflage, München, 1999.
- [Sal05] Saleck, Theo: Chefsache Open Source, 1. Auflage, Wiesbaden, 2005.
- [Sal10] Sallam, Rita; Hostmann, Bill; Richardson, James; Bitterer, Andreas: Magic Quadrant for Business Intelligence Platforms, 2010, URL: <http://www.gartner.com/technology/media-products/reprints/tableau/article1/article1.html>, (Abgerufen am: 18.03.2010).
- [Sch07] Schneider, Martin: Crossmedia-management, 1. Auflage, Köln, 2007.
- [Sch10] Schubert, Victoria: Solarindustrie verdoppelt Gewinnmargen auf 30 Prozent,

- 2006, URL: <http://www.presstext.de/news/060803025/solarindustrie-verdoppelt-gewinnmargen-auf-30-prozent/>, (Abgerufen am: 11.04.2010).
- [Sch00] Schütte, Reinhard; Rotthowe, Thomas; Holten, Roland: Data Warehouse Managementhandbuch. Konzepte, Software, Erfahrungen, 1. Auflage, Berlin, Heidelberg, 2000.
- [STA10] STAS GmbH: STAS GmbH - Unternehmen planen und steuern mit STAS CONTROL - Controlling, Reporting und Planung, 2010, URL: <http://www.stas.de>, (Abgerufen am: 01.04.2010).
- [Str08] Strohmeier, Stefan: Informationssysteme im Personalmanagement. Architektur-Funktionalität-Anwendung, 1. Auflage, Wiesbaden, 2008.
- [Tal10] Talend GmbH: Interview mit Deborah Brook, CEO WDCi, 2009, URL: http://www.talend.com/news/newsletter17_de_interview.php, (Abgerufen am: 02.03.2010).
- [Tec10] TechTarget: business intelligence, 2010, URL: http://searchdatamanagement.techtarget.com/sDefinition/0,,sid91_gci213571,00.html, (Abgerufen am: 27.02.2010).
- [Teg05] Tegel, Tatjana: Multidimensionale Konzepte zur Controllingunterstützung in kleinen und mittleren Unternehmen, 1. Auflage, Wiesbaden, 2005.
- [Tei08] Teich, Irene: Kolbenschlag, Walter; Reiners, Wilfried: Der richtige Weg zur Softwareauswahl. Lastenheft, Pflichtenheft, Compliance, Erfolgskontrolle., 1. Auflage, Görlitz, 2008.
- [Tot00] Totok, Andreas: Modellierung von OLAP- und Data-Warehouse-Systemen, 1. Auflage, Wiesbaden, 2000.
- [Vol09] Volitich, Dan: IBM Cognos 8. Business Intelligence., 1. Auflage, Heidelberg, München, Landsberg, Frechen, Hamburg, 2009.
- [von02] von Maur, Eitel: Winter, Robert (Hrsg.): Vom Data Warehouse zum Corporate Knowledge Center, 1. Auflage, St. Gallen, 2002.
- [Wil04] Wilmes, Caroline; Dietl, Helmut ; van der Velden, Remco : Die strategische Resource 'Data Warehouse'. Eine ressourcentheoretisch empirische Analyse, 1. Auflage, Wiesbaden, 2004.

- [Win10] Winter, Robert; Aier, Stephan : Informationssystem-Architektur ,2009, URL:
<http://www.enzyklopaedie-der-wirtschaftsinformatik.de/wi-enzyklopaedie/lexikon/daten-wissen/Informationsmanagement/Information-/Informationssystem-Architektur>,
(Abgerufen am: 01.02.2010).
- [Zim09] Zimmermann, Mark: Business Intelligence (Teil 1): Erster Einstieg,2009, URL:
http://www.tecchannel.de/server/sql/1738998/business_intelligence_teil_1_erster_einstieg/index7.html, (Abgerufen am: 12.11.2009).

IV Anlage A

Mit dem Aufkommen der ersten Dialog- und Transaktionssysteme in den 1960er Jahren wurden erstmals in den Unternehmen Daten in elektronischer Form generiert und abgespeichert. Während vor der Nutzung von Dialog- und Transaktionssystemen ein Mangel an Informationen bestand, sahen sich die Führungskräfte nach der Einführung dieser Systeme mit einer Informationsflut konfrontiert. Um die Manager mittels Datenverarbeitungswerkzeugen bei ihren Aufgaben zu unterstützen, mussten die Daten zu Planungs- und Kontrollzwecken entsprechend aufbereitet werden

Um die Aufbereitung der Daten zu ermöglichen, wurden Managementinformationssysteme entwickelt. MIS nutzten eine direkte Anbindung zur operativen Datenbasis und hatten das Ziel, Führungsinformationen in Echtzeit¹ aus den operativen Daten zu generieren und dem Management bedarfsgerecht bereitzustellen. Diese Anforderungen konnten aber aufgrund der unausgereiften technischen Möglichkeiten nur teilweise erfüllt werden, so dass bereits in den 1970er Jahren eine ablehnende Haltung gegenüber diesen Systemen zu verzeichnen war. Zu dieser Ablehnung führte auch die Unzufriedenheit der Anwender mit den präsentierten Ergebnissen, welche in der Regel in Form von Computerausdrucken vorlagen. Hinzu kam, dass sich die Daten nach wie vor grundlegend an der operativen Datenstruktur orientierten und daher nur bedingt für die Unterstützung des Managements geeignet waren.

Nach einer Phase der Frustration und Ernüchterung bezüglich der bescheidenen Auswertungsmöglichkeiten der MIS wurde mit einem neuen Ansatz in den 1970er Jahren versucht, deren Schwachpunkte zu beheben. Entscheidungsunterstützungssysteme (EUS) bzw. Decision Support Systeme (DSS) verfolgten das Ziel, das Verhalten von Fach- und Führungskräften bei der Lösung von Fachproblemen mit Hilfe von Modellen, Methoden und problembezogenen Daten abzubilden. Stand bei den MIS noch die reine Datenversorgung der Führungskräfte im Vordergrund, konzentrierten sich die DSS darauf, die Manager im Planungs- und Entscheidungsprozess zu unterstützen und damit die Qualität der Entscheidungen spürbar zu verbessern.²

1 Echtzeit bedeutet in diesem Zusammenhang, dass die Informationen im Moment der Abfrage aus den operativen Daten abgeleitet werden.

2 Vgl. [Glu08, S. 63]

Die Verwendung der DSS konzentrierte sich anfangs in Stabstellen und separaten Fachabteilungen, so dass sich die Systeme als dezentrale Lösung für spezielle Probleme etablierten. Durch diese lokale Ausrichtung gestaltete sich eine Integration in die unternehmensweite EDV schwierig und verhinderte so die Entwicklung von unternehmensübergreifenden Modellen zur Simultanplanung. Die Stärken von Entscheidungsunterstützungssystemen sind vor allem bei „erkanntem Problemlösungsbedarf und eingegrenztem Problemumfang“¹ zu finden, so dass sie sich bis heute in Unternehmen flächendeckend im Einsatz befinden.

In den 1980er Jahren wurde mit dem Erscheinen von anwendungsfreundlichen Benutzeroberflächen und der zunehmenden Vernetzung der vorhandenen Datenverarbeitungssysteme die Basis für die Weiterentwicklung der Managementinformationssysteme² zu Executive Information Systems (EIS), Chefinformationssystemen (CIS) bzw. Führungsinformationssystemen (FIS) geschaffen. Diese leistungsstarken Systeme waren aufgrund der technischen Entwicklung in der Lage, die Führungskräfte mit entscheidungsrelevanten Informationen zu versorgen und boten darüber hinaus erstmals intuitiv benutzbare und individuell anpassbare Oberflächen an. Die EIS waren in der Einführungsphase auf die Bedürfnisse des Top-Managements zugeschnitten und zeichneten sich durch eine Reihe von Innovationen wie der Einbindung mehrerer Datenquellen³ und heutigen Standardtechniken wie Drill-Down⁴ und Exception Reporting⁵ aus.

Zusammenfassen lassen sich die vorgestellten Managementinformationssysteme, Decision Support Systeme und Executive Information Systeme unter dem Begriff Management Support Systeme (MSS), der „alle Spielarten der elektronischen Unterstützung betrieblicher Entscheidungsträger bei der Abwicklung anfallender Aufgaben“⁶ umfasst.

1 [Cha06, S. 7]

2 Weiterentwicklung der MIS aus den 1960er Jahren

3 Als Datenquelle für ein EIS können die eigene EIS-Datenbasis, die operative Datenbasis, ein Data Warehouse oder auch externe Datenbestände dienen.

4 Disaggregierende Informationsanalyse

5 Berichtswesen mit Kennzeichnung auffälliger Abweichungen

6 [Cha06, S. 9]

V Anlage B

Artikel der Gartner Inc. zur Definition und Abgrenzung des Begriffs Business Intelligence Plattform

Market Definition/Description

BI platforms enable users to build applications that help organizations learn, understand and optimize their business. Gartner defines a BI platform as a software platform that delivers the 13 capabilities listed below. These capabilities are organized into three categories of functionality: integration, information delivery and analysis. In 2009, enhancing integration between BI platform components has been a major focus of megavendors digesting their numerous acquisitions. Information delivery continues to be a core focus of most BI projects today, but we see an increasing demand for tools that enable easier and more intuitive analysis to discover new insights. The Gartner definition of "BI platform" has remained mostly consistent from previous years, but we have added a 13th capability this year for search-based BI.

Integration:

- BI infrastructure — All tools in the platform should use the same security, metadata, administration, portal integration, object model and query engine, and should share the same look and feel.
- Metadata management — Not only should all tools leverage the same metadata, but search (Google-like) interface.
- Analysis:
 - OLAP — This enables end users to analyze data with extremely fast query and calculation performance, enabling a style of analysis known as "slicing and dicing." This capability could span a variety of storage architectures, such as relational, multidimensional and in-memory.
 - Interactive visualization — This gives the ability to display numerous aspects of the data more efficiently by using interactive pictures and charts, instead of rows and columns. Over time, advanced visualization will go beyond just slicing and dicing

data to include more process-driven BI projects, allowing all stakeholders to better understand the workflow through a visual representation.

- Predictive modeling and data mining — This capability enables organizations to classify categorical variables and to estimate continuous variables using advanced mathematical techniques. BI developers are able to integrate models easily into BI reports, dashboards and analysis.
- Scorecards — These take the metrics displayed in a dashboard a step further by applying them to a strategy map that aligns key performance indicators with a strategic objective. Scorecard metrics should be linked to related reports and information to perform further analysis. A scorecard implies the use of a performance management methodology such as Six Sigma or a balanced scorecard framework. the offering should provide a robust way to search, capture, store, reuse and publish metadata objects such as dimensions, hierarchies, measures, performance metrics and report layout objects.
- Development tools — The BI platform should provide a set of programmatic development tools and a visual development environment, coupled with a software developer's kit for creating BI applications, for integrating them into a business process and/or embedding them in another application. The BI platform should also enable developers to build BI applications without coding by using wizard-like components for a graphical assembly process. The development environment should also support Web services in performing common tasks such as scheduling, delivering, administering and managing. In addition, the BI application should assign and track events or tasks allotted to specific users, based on predefined business rules. Often, this capability is delivered by integrating with a separate portal or workflow tool.
- Collaboration — This capability enables BI users to share and discuss information and/or manage hierarchies and metrics via discussion threads, chat and annotations either embedded in the application or through integration with collaboration, analytical master data management (MDM) and social software.

Information delivery:

- **Reporting** — Reporting provides the ability to create formatted and interactive reports (parameterized) with highly scalable distribution and scheduling capabilities. In addition, BI platform vendors should handle a wide array of reporting styles (for example, financial, operational and performance dashboards) and should enable users to access and fully interact with BI content delivered to mobile devices.
- **Dashboards** — This subset of reporting includes the ability to publish formal, Web-based reports with intuitive interactive displays of information, including dials, gauges, sliders, check boxes and traffic lights. These displays indicate the state of the performance metric compared with a goal or target value. Increasingly, dashboards are used to disseminate real-time data from operational applications.
- **Ad hoc query** — This capability enables users to ask their own questions of the data, without relying on IT to create a report. In particular, the tools must have a robust semantic layer to allow users to navigate available data sources. These tools should include a disconnected analysis capability that enables users to access BI content and analyze data remotely without being connected to a server-based BI application. In addition, these tools should offer query governance and auditing capabilities to ensure that queries perform well.
- **Microsoft Office integration** — In some cases, BI platforms are used as a middle tier to manage, secure and execute BI tasks, but Microsoft Office (particularly Excel) acts as the BI client. In these cases, it is vital that the BI vendor provides integration with Microsoft Office, including support for document formats, formulas, data "refresh" and pivot tables. Advanced integration includes cell locking and write-back.
- **Search-based BI** — Applies a search index to both structured and unstructured data sources and maps them into a classification structure of dimensions and measures (often leveraging the BI semantic layer) that users can easily navigate and explore using a search (Google-like) interface.

Analysis:

- OLAP — This enables end users to analyze data with extremely fast query and calculation performance, enabling a style of analysis known as "slicing and dicing." This capability could span a variety of storage architectures, such as relational, multidimensional and in-memory.
- Interactive visualization — This gives the ability to display numerous aspects of the data more efficiently by using interactive pictures and charts, instead of rows and columns. Over time, advanced visualization will go beyond just slicing and dicing data to include more process-driven BI projects, allowing all stakeholders to better understand the workflow through a visual representation.
- Predictive modeling and data mining — This capability enables organizations to classify categorical variables and to estimate continuous variables using advanced mathematical techniques. BI developers are able to integrate models easily into BI reports, dashboards and analysis.
- Scorecards — These take the metrics displayed in a dashboard a step further by applying them to a strategy map that aligns key performance indicators with a strategic objective. Scorecard metrics should be linked to related reports and information to perform further analysis. A scorecard implies the use of a performance management methodology such as Six Sigma or a balanced scorecard framework.

VI Anlage C

Für Tabellen mit inhaltlichen Schnittmengen werden in den operativen Systemen unterschiedliche Zugriffsschlüssel verwendet. Daher werden bei direkter Zusammenführung von Tabellen zwangsläufig Schlüsselverletzungen auftreten. Zur Verhinderung von Schlüsseldisharmonien werden bei der Tabellenvereinigung unter Einbeziehung von Zuordnungstabellen aus den Primärschlüsseln der operativen Systeme eindeutige Schlüssel generiert, so dass übergreifende Auswertungen möglich sind. Um beispielsweise einen eindeutigen Artikelschlüssel zu erstellen, ist es möglich, Mandantenfelder, Buchungskreis, Kunden und Artikelnummer zusammenzuführen.

Des Weiteren gilt es, Datensätze von unterschiedlichen Tabellen, die über einen identischen Attributnamen bzw. über eine identische Bedeutung verfügen, jedoch durch unterschiedliche Wertebereiche oder Domänen gekennzeichnet werden, zu harmonisieren. Mit Hilfe einer Mapping Tabelle und der Bestimmung einer eindeutigen Domäne werden die unterschiedlich codierten Daten vereinheitlicht. So können zum Beispiel für die zwei geschlechtlichen Ausprägungen die Domänen 0 und 1 oder W und M in den operativen Systemen vorhanden sein, die es gilt, in eine einheitliche Form zu überführen. Dabei kann eine der bestehenden Domänen ausgewählt oder aber diese innerhalb der Mapping Tabellen einer neuen Domäne zugeordnet werden.

Als problematisch können sich auch Attribute erweisen, die unterschiedliche Bezeichnung aufweisen, inhaltlich jedoch äquivalent sind, sogenannte Synonyme. Dies können etwa unterschiedliche Attribute für einen Artikel sein, der in einer anderen Tabelle als Produkt bezeichnet wird. Zur Lösung führt die Festlegung einer einheitlichen Attributbezeichnung und die Anpassung der verbleibenden Attribute.

Beispielhaft für die Harmonisierung von Homonymen steht das Attribut Partner. Dieses Attribut kann auf einen Kunden oder auf einen Lieferant verweisen. Durch die Wahl und Vergabe von eindeutigen Attributnamen kann eine eindeutige Zuordnung der Partner erfolgen.

VII Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Alle Teile, die wörtlich oder sinngemäß einer Veröffentlichung entstammen, sind als solche kenntlich gemacht. Die Arbeit wurde noch nicht veröffentlicht oder einer anderen Prüfungsbehörde vorgelegt.

Mittweida, 29. Juni 2010

Thomas Gruschwitz